(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 July 2006 (27.07.2006)

PCT

(10) International Publication Number
WO 2006/079020 A2

(51) International Patent Classification:
C12N 9/78 (2006.01)

(21) International Application Number:
PCT/US2006/002265

(22) International Filing Date: 19 January 2006 (19.01.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/645,304 19 January 2005 (19.01.2005) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION** [US/US]; A144, Astecc Building, Lexington, Kentucky 40506-0286 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHAPPELL, Joe** [US/US]; 343 Plant Science Building, Plant & Soil Sciences, University Of Kentucky, Lexington, Kentucky 40546-0312 (US). **YEO, Yunsoo** [KR/KR]; National Institute of Agricultural Sciences and Technology, Division of Molecular Genetics, Rural Development Administration (RDA), Suwon 441-707 (KR). **TAKAHASHJI, Shunji** [US/US]; 343 Plant Science Building, 1405 Veterans Drive, University Of Kentucky, Lexington, Kentucky 40546-0312 (US).

(74) Agent: **FARER, Michael, B.**; CATALYST LAW GROUP, APC, 9710 Scranton Road, Suite 170, San Diego, CA 92121 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: FUNCTIONAL IDENTIFICATION OF THE HYOSCYAMUS MUTICUS GENE CODING FOR PREMNASPIRODIENE HYDROXYLASE ACTIVITY

(57) Abstract: Isolated and purified nucleic acid molecules encoding *Hyoscyamus muticus* premnaspirodiene oxidase (HPO) have been prepared by gene cloning. These nucleic acid molecules encode the oxidase enzyme which can catalyze the hydroxylation of valencene to β-nootkatol and nootkatone, which is industrially important and used in food products. Also within the scope of the invention are vectors including the nucleic acid molecules, host cells transformed or transfected by the vectors, and methods of producing proteins by expression of the nucleic acid molecules in the host cells. Further within the scope of the invention are methods of producing an oxidized terpene substrate from an unoxidized terpene substrate using proteins or expressed nucleic acid molecules according to the invention.

WO 2006/079020 A2

Functional Identification of the Hyoscyamus muticus Gene Coding for Premnaspirodiene Hydroxylase Activity

RELATED APPLICATION

[0001] Benefit of priority under 35 U.S.C. 119(e) is claimed herein to U.S. Provisional Application No.: 60/645,304, filed January 19, 2005. The disclosure of the above referenced application is incorporated by reference in its entirety herein.

FIELD OF THE INVENTION

[0002] The present invention relates to the functional identification of a gene which encodes for an enzyme that catalyzes the regio-and stereo-specific hydroxylation of sesquiterpene scaffolds. More particularly, the present invention discloses the DNA sequence for the *Hyoscyamus muticus* premnaspirodiene synthase gene, HPO, which when expressed in a heterologous host such as yeast, provides an enzyme activity that catalyzes mono-and successive hydroxylation of premnaspirodiene and other sesquiterpene scaffolds.

BACKGROUND OF THE INVENTION

[0003] Terpenes are a diverse family of compounds with carbon skeletons composed of five-carbon isoprene units. Approximately 20,000 different terpenes and terpenoids (compounds of terpene origin whose carbon skeleton has been altered or rearranged) have been identified to date, representing only a small fraction of the estimated natural variation. Terpenes are commonly isolated from the essential oils of plants. Essential oils often have pleasant tastes or aromas, and they are widely used as flavorings, deodorants, and medicines.

[0004] Sesquiterpenes are terpenes with 15 carbon atoms (three isoprene units). The plant kingdom contains the highest diversity of sesquiterpenes. Often they play a role in defense of the plants against pathogens, insects and herbivores and for attraction of pollinating insects.

[0005] Valencene (1,2,3,5,6,7,8,8a-octahydro-7-isopropenyl-1,8a-dimethyl-naphthalene) and nootkatone (4,4a,5,6,7,8-hexahydro-6-isopropenyl-4,4a-dimethyl-2(3H)-naphtalenone) are just two examples of sesquiterpenes that are derived from cyclization of the ubiquitous pyrophosphate intermediate farnesyl diphosphate. Nootkatone is formed by the oxidation of valencene.

[0006] Valencene and nootkatone are compounds of natural origin, and are natural constituents of citrus oils, such as orange and grapefruit. Because of its excellent organoleptic qualities and in particular its typical grapefruit taste, nootkatone is a widely used ingredient in perfumery and the flavor industry. Alternatively, nootkatone may be used as an insecticide. Valencene, the starting material for the generation of nootkatone (either biologically or chemically), is also used as a flavorant and fragrance.

[0007] Nootkatone is a high demand, high value flavorant added to many of the commercial soft drinks sold worldwide. Currently, the practice of extracting nootkatone from citrus pulp and rind is considered an expensive and somewhat unreliable process. Nootkatone can be synthesized by the oxidation of valencene. The valencene starting material is expensive and is easily degraded during evaporative heat concentration processes typically used to remove the bulk of water from the feed juice. Thus, current methods to purify valencene from citrus fruits are costly, difficult, and are limited by what the fruit can deliver. Moreover, such methods are vulnerable to interruptions in the supply of citrus fruits, which is dependent on the weather. A frost or hailstorm in a major citrus fruit growing region such as Florida can interrupt the supply. Furthermore, methods to produce nootkatone that consume valencene are quite costly, and thus not commercially desirable. Therefore, there is a need for an alternative means for preparing valencene and nootkatone.

SUMMARY OF THE INVENTION

[0008] Successive hydroxylation of valencene at C2 yields nootkatone, a high demand, high value flavor additive routinely used in the fragrance and food industries. The HPO gene of the present invention now provides an alternative means of generating important stereochemically pure starting materials for the reliable and cost effective production of nootkatone and other high value sesquiterpenes. For example, the co-expression of a valencene synthase gene along with the HPO gene in transgenic plants or microbial cells provides for large scale production of nootkatone. Alternatively, other terpene synthase genes can be used in combination with the HPO gene to generate other novel terpene moieties, which could be of value for pharmaceutical, agricultural and other industrial applications.

[0009] In one embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0010] In another embodiment, the present invention relates to an isolated and purified DNA sequence wherein the sequence is:

ATGCAATTCTTCAGCTTGGTTTCCATCTTCCTTTTTCTATCTTTTTTGTTTTTG
TTAAGGAAATGGAAGAACTCCAATAGCCAGTCCAAGAAATTGCCTCCAGGT
CCATGGAAACTTCCATTACTAGGTAGCATGCTTCATATGGTTGGTGGACTTC
CACATCATGTACTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCATCTT
CAACTTGGTGAAGTTTCTGCTGTTGTTGTTACTTCTCCTGATATGGCAAAAG
AAGTACTAAAAACTCATGACATTGCGTTCGCGTCTAGGCCTAAACTTTTAGC
CCCAGAGATTGTATGTTACAACAGGTCTGACATTGCGTTTTGCCCTTATGGT
GATTACTGGAGACAAATGCGTAAAATTTGTGTCTTGGAAGTGTTGAGTGCCA
AGAATGTTAGGTCATTCAGCTCTATTAGGCGCGATGAAGTGCTTCGTCTAGT
TAATTTTGTCCGATCATCTACGAGTGAGCCGGTTAACTTTACTGAAAGGCTG
TTTTTATTCACAAGTTCCATGACATGTAGATCAGCATTTGGGAAAGTGTTCAA

GGAACAGGAAACATTTATACAACTAATCAAAGAAGTGATAGGTTTAGCAGGA
GGATTTGATGTGGCTGACATCTTCCCATCACTGAAGTTCCTCCATGTACTAA
CTGGAATGGAGGGTAAGATTATGAAGGCTCACCATAAAGTAGATGCAATTGT
TGAGGATGTCATCAATGAGCATAAGAAGAACCTTGCAATGGGGAAAACTAAT
GGTGCATTAGGAGGTGAAGATCTAATTGATGTTCTTTTAAGACTTATGAATG
ATGGAGGCCTTCAATTTCCTATCACCAATGACAACATCAAAGCTATTATCTTT
GACATGTTTGCTGCTGGGACAGAGACTTCATCGTCAACACTTGTATGGGCTA
TGGTGCAAATGATGAGAAATCCAACTATACTAGCCAAAGCTCAAGCAGAAGT
AAGAGAAGCATTCAAAGGAAAAGAAACTTTCGATGAAAATGATGTCGAAGAG
TTGAAATACTTGAAGTTAGTCATTAAAGAAACTCTAAGACTCCATCCACCAGT
TCCACTTTTGGTCCCAAGAGAATGTAGGGAAGAAACAGAAATAAATGGCTAC
ACTATTCCAGTAAAGACCAAAGTCATGGTTAATGTTTGGGCATTAGGAAGAG
ATCCGAAATACTGGGACGACGCAGATAACTTCAAGCCAGAGAGATTTGAGC
AGTGTTCTGTGGACTTTATAGGTAACAATTTTGAATATCTTCCATTTGGTGGT
GGAAGGAGGATATGTCCAGGGATATCATTTGGTTTAGCTAATGTTTATTTGC
CATTGGCTCAATTGCTATATCATTTTGATTGGAAACTCCCTACTGGAATGGAA
CCAAAAGACTTGGATTTGACAGAATTGGTTGGAGTAACTGCTGCCAGAAAG
AGTGATCTTATGTTGGTTGCGACTCCTTATCAACCTTCTCGAGAGTAA (SEQ
ID NO: 1).

[0011] In yet another embodiment, the present invention relates to an isolated and purified DNA sequence wherein the sequence encodes a protein of the sequence:

MQFFSLVSIFLFLSFLFLLRKWKNSNSQSKKLPPGPWKLPLLGSMLHMVGGLPH
HVLRDLAKKYGPLMHLQLGEVSAVVVTSPDMAKEVLKTHDIAFASRPKLLAPEIV
CYNRSDIAFCPYGDYWRQMRKICVLEVLSAKNVRSFSSIRRDEVLRLVNFVRSS
TSEPVNFTERLFLFTSSMTCRSAFGKVFKEQETFIQLIKEVIGLAGGFDVADIFPS
LKFLHVLTGMEGKIMKAHHKVDAIVEDVINEHKKNLAMGKTNGALGGEDLIDVLL
RLMNDGGLQFPITNDNIKAIIFDMFAAGTETSSSTLVWAMVQMMRNPTILAKAQ
AEVREAFKGKETFDENDVEELKYLKLVIKETLRLHPPVPLLVPRECREETEINGY
TIPVKTKVMVNVWALGRDPKYWDDADNFKPERFEQCSVDFIGNNFEYLPFGG

GRRICPGISFGLANVYLPLAQLLYHFDWKLPTGMEPKDLDLTELVGVTAARKSD LMLVATPYQPSRE (SEQ ID NO: 2).

[0012] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 95% identical to SEQ ID NO: 1, provided that the nucleic acid sequence is translated into a protein encoding a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0013] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 97.5% identical to SEQ ID NO: 1.

[0014] In yet another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 99% identical to SEQ ID NO: 1.

[0015] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu

or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0016] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2.

[0017] In yet another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0018] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0019] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

In yet another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

[0020] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the

sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

[0021] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

[0022] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0023] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is SEQ ID NO: 1.

In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence encodes a protein of the sequence SEQ ID NO: 2.

In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 97.5% identical to SEQ ID NO: 1.

[0024] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 99% identical to SEQ ID NO: 1.

[0025] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0026] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to ten conservative amino acid substitutions from SEQ ID NO: 2.

[0027] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0028] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0029] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0030] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

[0031] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

[0032] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

[0033] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0034] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is SEQ ID NO: 1.

[0035] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence encodes a protein of the sequence SEQ ID NO: 2.

[0036] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 97.5% identical to SEQ ID NO: 1.

[0037] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 99% identical to SEQ ID NO: 1.

[0038] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0039] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to ten conservative amino acid substitutions from SEQ ID NO: 2.

[0040] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0041] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0042] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0043] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

[0044] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

[0045] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

[0046] In another embodiment, the present invention relates to a method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

[0047] In yet another embodiment, the present invention relates to a method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing a host cell according to the present invention under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

[0048] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0049] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein has the amino acid sequence of SEQ ID NO: 2.

[0050] In yet another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0051] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0052] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity

of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 5 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0053] In yet another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0054] In another embodiment, the present invention relates to an isolated and purified protein molecule having a functional terpene hydroxylase activity wherein the protein molecule has the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0055] In another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone.

[0056] In yet another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule of SEQ ID NO: 2 with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone.

[0057] In another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone, wherein the isolated and purified protein molecule has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0058] In another embodiment, the present invention relates to a yeast cell transformed or transfected with:

(a) a first vector including therein a DNA molecule encoding

functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone;

(b) a second vector including therein a DNA molecule encoding a functional P450 reductase protein; and

(c) a third vector including therein a DNA molecule encoding a functional *Hyoscyamus muticus* premnaspirodiene protein synthase;

such that the yeast cell expresses: (1) the functional *Hyoscyamus muticus* premnaspirodiene oxidase protein in a quantity sufficient to hydroxylate valencene; (2) the functional P450 reductase protein in a quantity sufficient to supply reducing equivalents for the *Hyoscyamus muticus* premnaspirodiene oxidase protein; and (3) the functional *Hyoscyamus muticus* premnaspirodiene protein synthase in a quantity sufficient to produce premnaspirodiene; and such that the premnaspirodiene is converted by the cell to solavetivone.

[0059] In yet another embodiment, the present invention relates to a method of producing an oxidized terpene from unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to a host cell transformed or transfected with a vector comprising isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0060] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence of SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0061] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA encoding a protein of SEQ ID NO: 2;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0062] In yet another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the DNA sequence is at least 95% identical to SEQ ID NO: 1 provided that the DNA sequence is translated into a protein encoding a functional *Hyoscyamus muticus* premnaspirodiene protein synthase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0063] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the sequence is at least 97.5% identical to SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0064] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the sequence is at least 99% identical to SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0065] In yet another embodiment, the present invention relates to a method for producing a mutein of *Hyoscyamus muticus* premnaspirodiene oxidase with at least one altered property selected from the group consisting of regiospecificity and stereospecificity comprising the steps of:

(a) threading the wild-type *Hyoscyamus muticus* premnaspirodiene oxidase amino acid sequence onto the three-dimensional structure of a mammalian analogue enzyme;

(b) docking at least one sesquiterpene substrate into the predicted active site pocket;

(c) mapping amino acid residues within a defined distance of the modeled substrate molecules and rationalizing the amino acid residues with

predicted chemical transformations catalyzed by the *Hyoscyamus muticus* premnaspirodiene oxidase to identify amino acid residues capable of steric, ionic, electronic and hydrophobic interactions with the substrate molecules;

(d) selecting at least one altered amino acid residue for its potential for substrate binding and positioning of the substrate molecule relative to the heme-catalytic center; and

(e) introducing the at least one altered amino acid amino acid residue by site-directed mutagenesis to produce the mutein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0066] The following invention will become better understood with reference to the specification, appended claims, and accompanying drawings, where:

[0067] Figure 1 illustrates biosynthetic transformations catalyzed by the Hyoscyamus muticus premnaspirodiene synthase, showing that the HPO enzyme catalyzes the successive hydroxylation of sesquiterpene scaffolds, thus generating a first mono-hydroxylated form, followed by the subsequent ketone form.

[0068] Figure 2 illustrates that the PCR primers which were used to isolate the HPO gene were initially designed relative to the EAH gene (71D20, GenBank accession number AF368376), and subsequently to other 71D family members whose expression are inducible by biological and abiotic stress (71D4, accession # AJ296346; 71D6, U48434; 71D7, U48435; 71D16, AF166332). Figure 3 illustrates the results of GC-MS analysis of reaction products generated by in vitro assays with microsomes from yeast over-expressing the HPO gene and incubated with premnaspirodiene as substrate (total ion chromatogram).

[0069] Figure 4 illustrates the results of GC-MS analysis of reaction

products generated by in vitro assays with microsomes from yeast over-expressing the HPO gene and incubated with valencene as substrate. Total ion chromatogram.

[0070] Figure 5 illustrates a comparison of mass spectra for reaction products generated in the in vitro assay (left panels) to authentic standards for ß-nootkatol and nootkatone (right panels), wherein RT=retention time.

[0071] Figure 6 shows that site-directed mutagenesis was used to alter the DNA sequence of the HPO gene corresponding to those codons for the indicated amino acids, the mutant genes over-expressed in yeast and isolated microsomes used to biochemically characterize the mutant enzyme activities.

[0072] Figure 7 provides a cartoon depiction of the engineering of a yeast line to effect sesquiterpene hydrocarbon and oxygenated terpene biosynthesis (A) and an ethyl acetate extract of P450 reductase, HPS and HPO engineered yeast cells assessed by GC-MS (B).

[0073] Figure 8 illustrates a sequence alignment of amino acids lining the active site (1st tier) and those within 3 Å of the active site residues (2nd tier) of HPH (SEQ ID NO: 2) with the corresponding positions of TEAH (SEQ ID NO: 19).

[0074] Figure 9 illustrates the mechanisms of reactions catalyzed by HPO relative to EAH, wherein the differing font size visually depicts relative catalytic rates.

[0075] Figure 10 illustrates the CO difference spectrum for HPO expression in yeast, wherein the absolute absorbance value can be used to calculate the absolute level of the HPO protein in the microsome preparation.

[0076] Figure 11 illustrates the reactions carried out with the substrate premnaspirodiene, indicating successive hydroxylation to 4β-solavetivol and then to solavetivone.

[0077] Figure 12 illustrates the reactions carried out with the substrate valencene, indicating successive hydroxylation to β-nootkatol and then to nootkatone.

[0078] Figure 13 illustrates the reaction carried out with valencene with a reaction time of 1 minute.

[0079] Figure 14 illustrates the reaction carried out with 5-epi-aristolochene with a reaction time of 0, 1, 2, 5 or 10 minutes.

[0080] Figure 15 illustrates the reaction carried out with 4-epieremophilene and its double-bond isomer with a reaction time of 5 minutes (upper panel), as well as the reaction with α-cedrene with a reaction time of 5 minutes (lower panel).

[0081] Figure 16 is a table showing a comparison of enzyme kinetics of HPO for various substrates relative to the previously characterized EAH (5-epi-aristolochene dihydroxylase) hydroxylases. K_m, k_{cat}, and k_{cat}/K_m are shown.

[0082] Figure 17 is a schematic depiction of a method for converting HPO into an enzyme with the same activity as EAH.

[0083] Figure 18 shows the EAH activity of a number of muteins, including those produced by domain swapping between EAH and HPO, and those produced by domain swapping between EAH and HPO with additional mutations in the HPO domain, including: V482I/A484I (a double mutation); V366S/V482I/A484I (a triple mutation); G280T/G281S/V366S/V482I/A484I (a

quintuple mutation); I294V/F296V/V366S/V482I/A484I (a quintuple mutation); I294V/F296V/V366S/V482I/A484I (a septuple mutation); G280T/G281S/V366S/V480S/V482I (a quintuple mutation); I294V/F296V/366S/V480S/V482I (a quintuple mutation; and I294V/F296V/V366S/V480S/A484I (a septuple mutation). Light bars indicate the EAH framework and dark bars indicate the HPO framework.

[0084] Figure 19 shows the general strategy of using domain-swapping mutations based on substrate recognition sequences (SRS), reciprocal site-directed mutagenesis based on homology modeling with mammalian P450s, and a combination of domain-swapping and site-directed mutagenesis. The six substrate recognition sites are shown along with EAH activity.

[0085] Figure 20 shows the results of homology modeling and site-directed mutagenesis in generating muteins of HPO indicated by the amino acid in the native (wildtype) enzyme, the amino acid position, then the mutant amino acid. For example, V366S indicates that the valine in position 366 of the wildtype enzyme has been mutated to serine. Mutants affecting SRS 4, 5, and 6 include: S308T/V366S/V480S/V482I/A484I; S308T/V366S/V480S/A484I; S308T/V366S/V482I/A484I; S308T/V366S/V480S/V482I/; S308T/V366S/A484I;S308T/V366S/V482I; and S308T/V366S/V480S. Mutants affecting SRS 4 and 5 include S308T/V366S. Mutants affecting SRS 5 and 6 include: V366S/V480S/V482I/A484I; V366S/V480S/A484I; V366S/V482I/A484I (designated the M3 mutant); V366S/V480S/V482I; V366S/A484I; and V366S/V480S. Mutants affecting SRS 5 include: V366S. Mutants affecting SRS6 include: V480S/V482I/A484I; V482I/A484I; V480S/A484I; V480S/V482I; A484I; V482I; and V480S. The 5-epiaristolochene hydroxylase activity of these enzymes is shown.

[0086] Figure 21 is similar to Figure 20, but depicts changes in SRS 1 and/or SRS 2 based on M3 (Figure 20). Mutants include: L52E/M3,

L52E/G209E/M3, G209E/M3, C119S/M3, R113Q/M3, V109E/M3, E107D/C119S/M3, E107D/M3, P106M/C119S/M3, P106M/M3, L104V/C119S/M3, L104V/P106M/E107D/M3, L104V/M3, L103I/L104V/C119S/M3, L103I/L104V/M3, L103I/C119S/M3, and L103I/M3. The 5-epiaristolochene hydroxylase activity of these enzymes is shown.

[0087] Figure 22 shows the results from a combination of domain-swapping mutations and site-directed mutagenesis; the mutation of V366S greatly diminishes the 5-epiaristolochene hydroxylase activity.

[0088] Figure 23 shows additional results from site-directed mutagenesis in SRS 4 as well as domain swapping. Mutants include: G280T/G281S/I294V/F296V/M3, I294V/F296V/M3, G280T/G281S/M3, V482I/A484I; V366S/V482I/A484I, and I294V/F296V/V366S/A482I/A484I. EAH activity is shown in terms of both 5-epiaristolochene hydroxylase activity and 1β(OH)EA hydroxylase activity.

DETAILED DESCRIPTION OF THE INVENTION

Abbreviations and Terms

[0089] In accordance with the present invention and as used herein, the following terms and abbreviations are defined with the following meanings, unless explicitly stated otherwise. These explanations are intended to be exemplary only. They are not intended to limit the terms as they are described or referred to throughout the specification. Rather, these explanations are meant to include any additional aspects and/or examples of the terms as described and claimed herein.

[0090] The following abbreviations are used herein:

[0091] The phrase "substantially identical" means that a relevant sequence is at least 70%, 75%, 80%, 85%, 90%, 92%, 95% 96%, 97%, 98%, or 99% identical to a given sequence. By way of example, such sequences may be allelic variants, sequences derived from various species, or they may be derived from the given sequence by truncation, deletion, amino acid substitution or addition. Percent identity between two sequences is determined by standard alignment algorithms such as ClustalX when the two sequences are in best alignment according to the alignment algorithm.

[0092] As used herein, the term "hybridization" or "hybridizes" under certain conditions is intended to describe conditions for hybridization and washes under which nucleotide sequences that are significantly identical or homologous to each other remain bound to each other. Appropriate hybridization conditions can be selected by those skilled in the art with minimal experimentation as exemplified in Ausubel, F. A., et al., eds., Current Protocols in Molecular Biology Vol. 2, John Wiley and Sons, Inc., New York (1995). Additionally, stringency conditions are described in Sambrook et al. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York (1989). Variations on the conditions for low, moderate, and high stringency are well known in the art and may be used with the current invention.

[0093] As used herein, the term "nucleic acid," "nucleic acid sequence," "polynucleotide," or similar terms, refers to a deoxyribonucleotide or ribonucleotide oligonucleotide or polynucleotide, including single- or double-stranded forms, and coding or non-coding (e.g., "antisense") forms. The term encompasses nucleic acids containing known analogues of natural nucleotides. The term also encompasses nucleic acids including modified or substituted bases as long as the modified or substituted bases interfere neither with the Watson-Crick binding of complementary nucleotides or with the binding of the nucleotide sequence by proteins that bind specifically. The term also encompasses nucleic-acid-like structures with synthetic backbones. DNA

backbone analogues provided by the invention include phosphodiester, phosphorothioate, phosphorodithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, morpholino carbamate, and peptide nucleic acids (PNAs); see Oligonucleotides and Analogues, a Practical Approach, edited by F. Eckstein, IRL Press at Oxford University Press (1991); Antisense Strategies, Annals of the New York Academy of Sciences, Volume 600, Eds. Baserga and Denhardt (NYAS 1992); Milligan (1993) J. Med. Chem. 36:1923-1937; Antisense Research and Applications (1993, CRC Press). PNAs contain non-ionic backbones, such as N-(2-aminoethyl) glycine units. Phosphorothioate linkages are described, e.g., by U.S. Pat. Nos. 6,031,092; 6,001,982; 5,684,148; see also, WO 97/03211; WO 96/39154; Mata (1997) Toxicol. Appl. Pharmacol. 144:189-197. Other synthetic backbones encompassed by the term include methylphosphonate linkages or alternating methylphosphonate and phosphodiester linkages (see, e.g., U.S. Pat. No. 5,962,674; Strauss-Soukup (1997) Biochemistry 36:8692-8698), and benzylphosphonate linkages (see, e.g., U.S. Pat. No. 5,532,226; Samstag (1996) Antisense Nucleic Acid Drug Dev 6:153-156). The nucleotide sequence or molecule may also be referred to as a "nucleotide probe." Some of the nucleic acid molecules of the invention are derived from DNA or RNA isolated at least once in substantially pure form and in a quantity or concentration enabling identification, manipulation, and recovery of its component nucleotide sequence by standard biochemical methods. Examples of such methods, including methods for PCR protocols that may be used herein, are disclosed in Sambrook et al. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York (1989), Ausubel, F. A., et al., eds., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., New York (1987), and Innis, M., et al. (Eds.) PCR Protocols: A Guide to Methods and Applications, Academic Press, San Diego, California (1990). Reference to a nucleic acid molecule also includes its complement as determined by the standard Watson-Crick base-pairing rules, with uracil (U) in RNA replacing thymine (T) in DNA, unless the complement is specifically excluded.

[0094] As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their well-known, three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used routinely in the art.

[0095] As described herein, the nucleic acid molecules of the invention include DNA in both single-stranded and double-stranded form, as well as the DNA or RNA complement thereof. DNA includes, for example, DNA, genomic DNA, chemically synthesized DNA, DNA amplified by PCR, and combinations thereof. Genomic DNA, including translated, non-translated and control regions, may be isolated by conventional techniques, e.g., using any one of the cDNAs of the invention, or suitable fragments thereof, as a probe, to identify a piece of genomic DNA which can then be cloned using methods commonly known in the art.

[0096] Polypeptides encoded by the nucleic acids of the invention are encompassed by the invention. As used herein, reference to a nucleic acid "encoding" a protein or polypeptide encompasses not only cDNAs and other intronless nucleic acids, but also DNAs, such as genomic DNA, with introns, on the assumption that the introns included have appropriate splice donor and acceptor sites that will ensure that the introns are spliced out of the corresponding transcript when the transcript is processed in a eukaryotic cell. Due to the degeneracy of the genetic code wherein more than one codon can encode the same amino acid, multiple DNA sequences can code for the same polypeptide. Such variant DNA sequences can result from genetic drift or artificial manipulation (e.g., occurring during PCR amplification or as the product of deliberate mutagenesis of a native sequence). Deliberate mutagenesis of a native sequence can be carried out using numerous techniques well known in the art. For example, oligonucleotide-directed site-specific mutagenesis procedures can be employed, particularly where it is desired to mutate a gene such that

predetermined restriction nucleotides or codons are altered by substitution, deletion or insertion. Exemplary methods of making such alterations are disclosed by Walder et al. (Gene 42:133,1986); Bauer et al. (Gene 37:73, 1985); Craik (BioTechniques, January 12-19, 1985); Smith et al. (Genetic Engineering: Principles and Methods, Plenum Press, 1981); Kunkel (Proc. Natl. Acad. Sci. USA 82:488,1985); Kunkel et al. (Methods in Enzymol. 154.367, 1987). The present invention thus encompasses any nucleic acid capable of encoding a protein of the current invention.

[0097] DNA sequences encoding the polypeptides or proteins of the invention can be obtained by several methods. For example, the DNA can be isolated using hybridization procedures that are well known in the art. These include, but are not limited to: (1) hybridization of probes to genomic or cDNA libraries to detect shared nucleotide sequences; (2) antibody screening of expression libraries to detect shared structural features; and (3) synthesis by the polymerase chain reaction (PCR). RNA sequences of the invention can be obtained by methods known in the art (See, for example, Current Protocols in Molecular Biology, Ausubel, et al., Eds., 1989).

[0098] The development of specific DNA sequences encoding proteins or polypeptides of the invention can be obtained by: (1) isolation of a double-stranded DNA sequence from the genomic DNA; (2) chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and (3) in vitro synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a eukaryotic donor cell. In the latter case, a double-stranded DNA complement of mRNA is eventually formed which is generally referred to as cDNA. Of these three methods for developing specific DNA sequences for use in recombinant procedures, the isolation of genomic DNA is the least common. This is especially true when it is desirable to obtain the microbial expression of eukaryotic polypeptides due to the presence of introns. For obtaining proteins or polypeptides according to the present

invention, the synthesis of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct synthesis of DNA sequences is not possible and the method of choice is the formation of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid-carrying cDNA libraries which are derived from reverse transcription of mRNA which is abundant in donor cells that have a high level of genetic expression. When used in combination with polymerase chain reaction technology, even rare expression products can be clones. In those cases where significant portions of the amino acid sequence of the polypeptide are known, the production of labeled single or double-stranded DNA or RNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single-stranded form (Jay, et al., Nucleic Acid Research 11:2325, 1983).

[0099] With respect to nucleotide sequences that are within the scope of the invention, all nucleotide sequences encoding the proteins or polypeptides that are embodiments of the invention as described are included in nucleotide sequences that are within the scope of the invention. This further includes all nucleotide sequences that encode polypeptides according to the invention that incorporate conservative amino acid substitutions as defined above. This further includes nucleotide sequences that encode larger proteins incorporating the proteins or polypeptides, including fusion proteins, and proteins that incorporate amino-terminal or carboxyl-terminal flanking sequences.

[0100] Nucleic acid sequences of the present invention further include nucleic acid sequences that are at least 95% identical to the sequences above, with the proviso that the nucleic acid sequences retain the activity of the sequences before substitutions of bases are made, including any activity of

proteins that are encoded by the nucleotide sequences and any activity of the nucleotide sequences that is expressed at the nucleic acid level, such as the binding sites for proteins affecting transcription. Preferably, the nucleic acid sequences are at least 97.5% identical. More preferably, they are at least 99% identical. For these purposes, "identity" is defined according to the Needleman-Wunsch algorithm (S.B. Needleman & C.D. Wunsch, "A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins," J. Mol. Biol. 48: 443-453 (1970)).

[0101] The current invention provides for isolated polypeptides. As used herein, the term "polypeptides" refers to a genus of polypeptide or peptide fragments that encompass the amino acid sequences identified herein, as well as smaller fragments. Alternatively, a polypeptide may be defined in terms of its antigenic relatedness to any peptide encoded by the nucleic acid sequences of the invention. Thus, in one embodiment, a polypeptide within the scope of the invention is defined as an amino acid sequence comprising a linear or 3-dimensional epitope shared with any peptide encoded by the nucleic acid sequences of the invention. Alternatively, a polypeptide within the scope of the invention is recognized by an antibody that specifically recognizes any peptide encoded by the nucleic acid sequences of the invention. Antibodies are defined to be specifically binding if they bind polypeptides of the invention with a K_a of greater than or equal to about 10^7 M^{-1}, such as greater than or equal to 10^8 M^{-1}. As used herein, the term "isolated," in reference to polypeptides or proteins, means that the polypeptide or protein is substantially removed from polypeptides, proteins, nucleic acids, or other macromolecules with which it, or its analogues, occurs in nature. Although the term "isolated" is not intended to require a specific degree of purity, typically, the protein will be at least about 75% pure, more typically at least about 90% pure, preferably at least about 95% pure, and more preferably at least about 99% pure.

[0102] A polypeptide "variant" as referred to herein means a polypeptide substantially homologous to a native polypeptide, but which has an amino acid sequence different from that encoded by any of the nucleic acid sequences of the invention because of one or more deletions, insertions or substitutions. Variants can comprise conservatively substituted sequences, meaning that a given amino acid residue is replaced by a residue having similar physiochemical characteristics. See Zubay, Biochemistry, Addison-Wesley Pub. Co., (1983). In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and may be made generally without altering the biological activity of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, e.g. Watson et al. Molecular Biology of the Gene, 4th Edition, 1987, Benjamin/Cummings, p. 224). In particular, such a conservative variant has a modified amino acid sequence, such that the change(s) do not substantially alter the protein's (the conservative variant's) structure and/or activity, e.g., antibody activity, enzymatic activity, or receptor activity. These include conservatively modified variations of an amino acid sequence, i.e., amino acid substitutions, additions or deletions of those residues that are not critical for protein activity, or substitution of amino acids with residues having similar properties (e.g., acidic, basic, positively or negatively charged, polar or non-polar, etc.) such that the substitutions of even critical amino acids does not substantially alter structure and/or activity. Conservative substitution tables providing functionally similar amino acids are well known in the art. For example, one exemplary guideline to select conservative substitutions includes (original residue followed by exemplary substitution): Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu. An alternative exemplary guideline uses the following six groups, each containing amino acids that are conservative substitutions for one another: (1) alanine (A or Ala), serine (S or Ser), threonine

(T or Thr); (2) aspartic acid (D or Asp), glutamic acid (E or Glu); (3) asparagine (N or Asn), glutamine (Q or Gln); (4) arginine (R or Arg), lysine (K or Lys); (5) isoleucine (I or Ile), leucine (L or Leu), methionine (M or Met), valine (V or Val); and (6) phenylalanine (F or Phe), tyrosine (Y or Tyr), tryptophan (W or Trp); (see also, e.g., Creighton (1984) Proteins, W. H. Freeman and Company; Schulz and Schimer (1979) Principles of Protein Structure, Springer-Verlag). One of skill in the art will appreciate that the above-identified substitutions are not the only possible conservative substitutions. For example, for some purposes, one may regard all charged amino acids as conservative substitutions for each other whether they are positive or negative. In addition, individual substitutions, deletions or additions that alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence can also be considered "conservatively modified variations" when the three-dimensional structure and the function of the protein to be delivered are conserved by such a variation. As further exemplified below in Example 5, it is generally preferred that such conservative amino acid substitutions be made outside of the substrate recognition sequences (SRS) SRS1, SRS2, SRS3, SRS4, SRS5, or SRS6. However in some cases these conservative amino acid substitutions can be made in the substrate recognition sequences SRS1, SRS2, SRS3, SRS4, SRS5, or SRS6 if molecular modeling according to techniques known in the art indicates that such substitutions do not alter the enzyme specificity.

[0103] The effects of such substitutions can be calculated using substitution score matrices such PAM120, PAM-200, and PAM-250 as discussed in Altschul, (J. Mol. Biol. 219:55565 (1991)). Other such conservative substitutions, for example, substitutions of entire regions having similar hydrophobicity characteristics, are well known.

[0104] Naturally-occurring peptide variants are also encompassed by the invention. Examples of such variants are proteins that result from alternate mRNA splicing events or from proteolytic cleavage of the polypeptides described

herein. Variations attributable to proteolysis include, for example, differences in the N- or C-termini upon expression in different types of host cells, due to proteolytic removal of one or more terminal amino acids from the polypeptides encoded by the sequences of the invention.

[0105] Variants of the valencene synthase of the invention may be used to attain desired enhanced or reduced enzymatic activity, modified regiochemistry or stereochemistry, or altered substrate utilization or product distribution. A variant or site direct mutant may be made by any methods known in the art. Variants and derivatives of native polypeptides can be obtained by isolating naturally-occurring variants, or the nucleotide sequence of variants, of other or same plant lines or species, or by artificially programming mutations of nucleotide sequences coding for native citrus polypeptides. Methods for site-directed mutagenesis are well-known in the art and include, but are not limited to, the methods described in J. Sambrook & D.W. Russell, "Molecular Cloning: A Laboratory Manual (3d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001), ch. 13, incorporated herein by this reference.

[0106] In one embodiment, the invention contemplates: vectors comprising the nucleic acids of the invention. As used herein, the term "vector" refers to a plasmid, virus, phagemid, or other vehicle known in the art that has been manipulated by insertion or incorporation of heterologous DNA, such as nucleic acid encoding proteins or polypeptides according to the present invention. Vectors include, but are not limited to, expression vectors; vectors suitable for other purposes that are not expression vectors are known in the art. Such expression vectors typically contain a promoter sequence for efficient transcription of the inserted nucleic acid in a cell. The expression vector typically contains an origin of replication, a promoter, as well as specific genes that permit phenotypic selection of transformed cells.

[0107] Expression vectors containing a nucleic acid sequence of the invention can be prepared using well known methods and include a cDNA sequence encoding the polypeptide operably linked to suitable transcriptional or translational regulatory nucleotide sequences. Examples of regulatory sequences include transcriptional promoters, operators, or enhancers, mRNA ribosomal binding sites, and appropriate sequences which control transcription and translation initiation and termination. Nucleotide sequences are "operably linked" when the regulatory sequence functionally relates to the cDNA sequence of the invention. Expression vectors, regulatory elements and the construction thereof are well known in the art, and therefore are not limited to those recited above.

[0108] In addition, sequences encoding appropriate signal peptides that are not naturally associated with the polypeptides of the invention can be incorporated into expression vectors. For example, a DNA sequence for a signal peptide (secretory) leader can be fused in-frame to a nucleotide sequence of the invention so that the polypeptide of the invention is initially translated as a fusion protein comprising the signal peptide. A signal peptide that is functional in the intended host cells enhances extracellular secretion of the expressed polypeptide. The signal peptide can be cleaved from the polypeptide upon secretion from the cell. In some cases, signal peptides are cleaved in two or more stages; this is also within the scope of the invention where appropriate.

[0109] Fusions of additional peptide sequences at the amino and carboxyl terminal ends of the polypeptides of the invention can be used with the current invention.

[0110] In one embodiment, the invention includes a host cell comprising a nucleic acid of the invention. As used herein, the term "host cell" refers to a cell in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny

may not be identical to the parental cell since there may be mutations that occur during replication. Such progeny are included when the term "host cell" is used. Methods of stable transfer where the foreign DNA is continuously maintained in the host are known in the art. Another embodiment of the invention is a method of making a recombinant host cell comprising introducing the vectors of the invention, into a host cell. In a further embodiment, a method of producing a polypeptide comprising culturing the host cells of the invention under conditions to produce the polypeptide is contemplated. In one embodiment the polypeptide is recovered.

[0111] Suitable host cells for expression of polypeptides of the invention are well known in the art, and include, but are not limited to, prokaryotes, yeast, higher eukaryotic cells, or combinations thereof. (See for example, Pouwels et al. Cloning Vectors: A Laboratory Manual, Elsevier, New York (1985)). A particularly preferred host cell is yeast. Cell-free translation systems, also well known in the art, could also be employed to produce the disclosed polypeptides using RNAs derived from DNA constructs disclosed herein.

[0112] Host cells may be modified by any methods known in the art for gene transfer including, for example, the use of delivery devices such as lipids and viral vectors, naked DNA, electroporation and particle-mediated gene transfer.

[0113] Transformation of a host cell with recombinant DNA may be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *Escherichia coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method by procedures well known in the art. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell or by electroporation.

[0114] When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used.

[0115] A variety of host-expression vector systems may be utilized to express proteins or polypeptides according to the present invention. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing an appropriate coding sequence; yeast transformed with recombinant yeast expression vectors containing the appropriate coding sequence; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing an appropriate coding sequence; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing an appropriate coding sequence; or animal cell systems infected with recombinant virus expression vectors (e.g., retroviruses, adenovirus, vaccinia virus) containing an appropriate coding sequence, or transformed animal cell systems engineered for stable expression. In such cases where glycosylation may be important, expression systems that provide for translational and post-translational modifications may be used; e.g., mammalian, insect, yeast or plant expression systems.

[0116] Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, etc. may be used in the expression vector (see e.g., Bitter, et al., Methods in Enzymology, 153:516-544, 1987). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage λ, plac, ptrp, ptac

(ptrp-lac hybrid promoter) and the like may be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus 7.5K promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted polypeptide coding sequence.

[0117] In yeast, such as *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters may be used. For a review see, Current Protocols in Molecular Biology, Vol. 2, 1988, Ed. Ausubel, et al., Greene Publish. Assoc. & Wiley Interscience, Ch. 13; Grant, et al., 1987, Expression and Secretion Vectors for Yeast, in Methods in Enzymology, Eds. Wu & Grossman, 31987, Acad. Press, N.Y., Vol. 153, pp.516-544; Glover, 1986, DNA Cloning, Vol. II, IRL Press, Wash., D.C., Ch. 3; and Bitter, 1987, Heterologous Gene Expression in Yeast, Methods in Enzymology, Eds. Berger & Kimmel, Acad. Press, N.Y., Vol. 152, pp. 673-684; and The Molecular Biology of the Yeast Saccharomyces, 1982, Eds. Strathern et al., Cold Spring Harbor Press, Vols. I and II. A constitutive yeast promoter such as ADH or LEU2 or an inducible promoter such as GAL may be used (Cloning in Yeast, Ch. 3, R. Rothstein In: DNA Cloning Vol. 11, A Practical Approach, Ed. DM Glover, 1986, IRL Press, Wash., D.C.). Alternatively, vectors may be used which promote integration of foreign DNA sequences into the yeast chromosome. Cloning and expression in yeast is further described in T.A. Brown, "Gene Cloning and DNA Analysis" (4^{th} ed., Blackwell, 2001), pp. 286-288. Other species of yeast such as *Pichia pastoris* can be used.

[0118] In cases where plant expression vectors are used, the expression of a polypeptide coding sequence may be driven by any of a number of promoters. For example, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV (Brisson, et al., Nature, 310:511-514, 1984), or the coat

protein promoter to TMV (Takamatsu, et al., EMBO J., 6:307-311, 1987) may be used; alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi, et al., EMBO J. 3:1671-1680, 1984; Broglie, et al., Science 224:838-843, 1984); or heat shock promoters, e.g., soybean hsp17.5-E or hsp17.3-B (Gurley, et al., Mol. Cell. Biol., 6:559-565, 1986) may be used. These constructs can be introduced into plant cells using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, microinjection, electroporation, etc. For reviews of such techniques see, for example, Weissbach & Weissbach, Methods for Plant Molecular Biology, Academic Press, NY, Section VIII, pp. 421-463, 1988; and Grierson & Corey, Plant Molecular Biology, 2d Ed., Blackie, London, Ch. 7-9, 1988.

[0119] An alternative expression system that can be used to express a protein of the invention is an insect system. In one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The polypeptide coding sequence may be cloned into non-essential regions (in *Spodoptera frugiperda*, for example, the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the polypeptide coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect cells in which the inserted gene is expressed. (E.g., see Smith, et al., J. Biol. 46:584, 1983; Smith, U.S. Pat. No. 4,215,051).

[0120] For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the a cDNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a

selectable marker. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., Cell 11:223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA, 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy, et al., Cell, 22:817, 1980) genes, which can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells respectively. Also, antimetabolite resistance-conferring genes can be used as the basis of selection; for example, the genes for dhfr, which confer resistance to methotrexate (Wigler, et al., Natl. Acad. Sci. USA,77:3567, 1980; O'Hare, et al., Proc. Natl. Acad. Sci. USA, 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci. USA, 78:2072, 1981; neo, which confers resistance to the aminoglycoside G418 (Colberre-Garapin, et al., J. Mol. Biol., 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre, et al., Gene, 30:147, 1984). Recently, additional selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, Proc. Natl. Acad. Sci. USA, 85:804, 1988); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue L., In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed., 1987).

[0121] Isolation and purification of protein expressed in mammalian, plant, yeast, or bacterial cells may be carried out by conventional means including preparative chromatography and immunological separations involving monoclonal or polyclonal antibodies. Antibodies provided in the present invention

are immunoreactive with a polypeptide according to the present invention. Antibody which consists essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations are provided. Monoclonal antibodies are made from antigen containing fragments of the protein by methods well known in the art (Kohler, et al., Nature, 256:495, 1975; Current Protocols in Molecular Biology, Ausubel, et al., ed., 1989).

[0122] In one embodiment, the cDNAs of the invention may be expressed in such a way as to produce either sense or antisense RNA. The expression of antisense RNA can be used to down-modulate the expression of the protein encoded by the mRNA to which the antisense RNA is complementary.

[0123] A further embodiment of the invention is methods of making terpenoids and sesquiterpene compounds, for example, using the nucleotides and polypeptides of the invention.

[0124] As used herein an acyclic pyrophosphate terpene precursor is any acyclic pryrophosphate compound that is a precursor to the production of at least one terpene including but not limited to geranyl-pyrophosphate (GPP), farnesyl-diphosphate (FPP) and geranylgeranyl-pyrophosphate (GGPP).

[0125] Also within the practice of the invention is an organism (e.g., microorganism or plant) that is used to construct a platform for high level production of a substrate of sesquiterpene synthases (e.g., FPP) and the introduction of a nucleic acid of the invention into the organism.

[0126] Unless otherwise indicated, nucleic acids of the invention that are DNA encompass both cDNA (DNA reverse transcribed from mRNA and lacking introns) and isolated genomic DNA (DNA that can contain introns.)

[0127] In one embodiment, the nucleic acids of the invention are used to create other nucleic acids coding for sesquiterpene synthases. For example, the invention provides for a method of identifying a sesquiterpene synthases comprising constructing a DNA library using the nucleic acids of the invention, screening the library for nucleic acids which encode for at least one sesquiterpene synthase. The DNA library using the nucleic acids of the invention may be constructed by any process known in the art where DNA sequences are created using the nucleic acids of the invention as a starting point, including but not limited to DNA shuffling. In such a method, the library may be screened for sesquiterpene synthases using a functional assay to find a target nucleic acid that encodes a sesquiterpene synthase. The activity of a sesquiterpene synthase may be analyzed using, for example, the methods described herein. In one embodiment, high through put screening is utilized to analyze the activity of the encoded polypeptides.

[0128] As used herein a "nucleotide probe" is defined as an oligonucleotide or polynucleotide capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, through complementary base pairing, or through hydrogen bond formation.

[0129] A "target nucleic acid" herein refers to a nucleic acid to which the nucleotide probe or molecule can specifically hybridize. The probe is designed to determine the presence or absence of the target nucleic acid, and the amount of target nucleic acid. The target nucleic acid has a sequence that is significantly complementary to the nucleic acid sequence of the corresponding probe directed to the target so that the probe and the target nucleic acid can hybridize. Preferably, the hybridization conditions are such that hybridization of the probe is specific for the target nucleic acid. As recognized by one of skill in the art, the probe may also contain additional nucleic acids or other moieties, such as labels, which may not specifically hybridize to the target. The term target nucleic acid may refer to the specific nucleotide sequence of a larger nucleic acid to which the

probe is directed or to the overall sequence (e.g., gene or mRNA). One skilled in the art will recognize the full utility under various conditions.

[0130] Other than in the operating example, or where otherwise indicated, all numbers expressing quantities of ingredients, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless indicated to the contrary, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should be construed in light of the number of significant digits and ordinary rounding approaches.

[0131] Notwithstanding that the numerical ranges and parameters setting forth the broad scope of the invention are approximations, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value; however, inherently contains certain errors necessarily resulting from the standard deviation found in their respective testing measurements.

[0132] In one embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone. The nucleic acid sequence is typically DNA. Typically, the DNA encodes a protein of the sequence
MQFFSLVSIFLFLSFLFLLRKWKNSNSQSKKLPPGPWKLPLLGSMLHMVGGLPH
HVLRDLAKKYGPLMHLQLGEVSAVVVTSPDMAKEVLKTHDIAFASRPKLLAPEIV
CYNRSDIAFCPYGDYWRQMRKICVLEVLSAKNVRSFSSIRRDEVLRLVNFVRSS
TSEPVNFTERLFLFTSSMTCRSAFGKVFKEQETFIQLIKEVIGLAGGFDVADIFPS
LKFLHVLTGMEGKIMKAHHKVDAIVEDVINEHKKNLAMGKTNGALGGEDLIDVLL

RLMNDGGLQFPITNDNIKAIIFDMFAAGTETSSSTLVWAMVQMMRNPTILAKAQ
AEVREAFKGKETFDENDVEELKYLKLVIKETLRLHPPVPLLVPRECREETEINGY
TIPVKTKVMVNVWALGRDPKYWDDADNFKPERFEQCSVDFIGNNFEYLPFGG
GRRICPGISFGLANVYLPLAQLLYHFDWKLPTGMEPKDLDLTELVGVTAARKSD
LMLVATPYQPSRE (SEQ ID NO: 2). However, the DNA sequence can alternatively encode a protein including SEQ ID NO: 2 therein, such as a protein that has additional amino-terminal or carboxyl-terminal flanking sequences, or a fusion protein coupling the amino acid sequence of SEQ ID NO: 2 with one or more additional domains. A particularly preferred DNA sequence is SEQ ID NO: 1, below.

[0133] In another embodiment, the present invention relates to an isolated and purified DNA sequence wherein the sequence is:
ATGCAATTCTTCAGCTTGGTTTCCATCTTCCTTTTTCTATCTTTTTTGTTTTTG
TTAAGGAAATGGAAGAACTCCAATAGCCAGTCCAAGAAATTGCCTCCAGGT
CCATGGAAACTTCCATTACTAGGTAGCATGCTTCATATGGTTGGTGGACTTC
CACATCATGTACTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCATCTT
CAACTTGGTGAAGTTTCTGCTGTTGTTGTTACTTCTCCTGATATGGCAAAAG
AAGTACTAAAAACTCATGACATTGCGTTCGCGTCTAGGCCTAAACTTTTAGC
CCCAGAGATTGTATGTTACAACAGGTCTGACATTGCGTTTTGCCCTTATGGT
GATTACTGGAGACAAATGCGTAAAATTTGTGTCTTGGAAGTGTTGAGTGCCA
AGAATGTTAGGTCATTCAGCTCTATTAGGCGCGATGAAGTGCTTCGTCTAGT
TAATTTTGTCCGATCATCTACGAGTGAGCCGGTTAACTTTACTGAAAGGCTG
TTTTTATTCACAAGTTCCATGACATGTAGATCAGCATTTGGGAAAGTGTTCAA
GGAACAGGAAACATTTATACAACTAATCAAAGAAGTGATAGGTTTAGCAGGA
GGATTTGATGTGGCTGACATCTTCCCATCACTGAAGTTCCTCCATGTACTAA
CTGGAATGGAGGGTAAGATTATGAAGGCTCACCATAAAGTAGATGCAATTGT
TGAGGATGTCATCAATGAGCATAAGAAGAACCTTGCAATGGGGAAAACTAAT
GGTGCATTAGGAGGTGAAGATCTAATTGATGTTCTTTTAAGACTTATGAATG
ATGGAGGCCTTCAATTTCCTATCACCAATGACAACATCAAAGCTATTATCTTT
GACATGTTTGCTGCTGGGACAGAGACTTCATCGTCAACACTTGTATGGGCTA

TGGTGCAAATGATGAGAAATCCAACTATACTAGCCAAAGCTCAAGCAGAAGT
AAGAGAAGCATTCAAAGGAAAAGAAACTTTCGATGAAAATGATGTCGAAGAG
TTGAAATACTTGAAGTTAGTCATTAAAGAAACTCTAAGACTCCATCCACCAGT
TCCACTTTTGGTCCCAAGAGAATGTAGGGAAGAAACAGAAATAAATGGCTAC
ACTATTCCAGTAAAGACCAAAGTCATGGTTAATGTTTGGGCATTAGGAAGAG
ATCCGAAATACTGGGACGACGCAGATAACTTCAAGCCAGAGAGATTTGAGC
AGTGTTCTGTGGACTTTATAGGTAACAATTTTGAATATCTTCCATTTGGTGGT
GGAAGGAGGATATGTCCAGGGATATCATTTGGTTTAGCTAATGTTTATTTGC
CATTGGCTCAATTGCTATATCATTTTGATTGGAAACTCCCTACTGGAATGGAA
CCAAAAGACTTGGATTTGACAGAATTGGTTGGAGTAACTGCTGCCAGAAAG
AGTGATCTTATGTTGGTTGCGACTCCTTATCAACCTTCTCGAGAGTAA (SEQ
ID NO: 1).

[0134] In yet another embodiment, the present invention relates to an isolated and purified DNA sequence wherein the sequence encodes a protein of the sequence:
MQFFSLVSIFLFLSFLFLLRKWKNSNSQSKKLPPGPWKLPLLGSMLHMVGGLPH
HVLRDLAKKYGPLMHLQLGEVSAVVVTSPDMAKEVLKTHDIAFASRPKLLAPEIV
CYNRSDIAFCPYGDYWRQMRKICVLEVLSAKNVRSFSSIRRDEVLRLVNFVRSS
TSEPVNFTERLFLFTSSMTCRSAFGKVFKEQETFIQLIKEVIGLAGGFDVADIFPS
LKFLHVLTGMEGKIMKAHHKVDAIVEDVINEHKKNLAMGKTNGALGGEDLIDVLL
RLMNDGGLQFPITNDNIKAIIFDMFAAGTETSSSTLVWAMVQMMRNPTILAKAQ
AEVREAFKGKETFDENDVEELKYLKLVIKETLRLHPPVPLLVPRECREETEINGY
TIPVKTKVMVNVWALGRDPKYWDDADNFKPERFEQCSVDFIGNNFEYLPFGG
GRRICPGISFGLANVYLPLAQLLYHFDWKLPTGMEPKDLDLTELVGVTAARKSD
LMLVATPYQPSRE (SEQ ID NO: 2).

[0135] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 95% identical to SEQ ID NO: 1, provided that the

nucleic acid sequence is translated into a protein encoding a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0136] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 97.5% identical to SEQ ID NO: 1.

[0137] In yet another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is a nucleic acid sequence that is at least 99% identical to SEQ ID NO: 1.

[0138] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0139] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2.

[0140] In yet another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0141] In another embodiment, the present invention relates to an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the DNA sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0142] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine, designated herein as V480S.

[0143] In yet another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the

sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine, designated herein as V482I.

[0144] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine, designated herein as A484I.

[0145] In another embodiment, the present invention relates to an isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine, designated herein as V366S.

[0146] In general, mutated proteins that are within the scope of the present invention are designated and described by reciting first the wild-type amino acid, the position at which the mutation occurs, then the mutated amino acid introduced, such as V366S, described above.

[0147] Mutated nucleic acid sequences and proteins encoded by the mutated nucleic acid sequences that are within the scope of the invention include multiple mutations that include more than one of the mutations described above or additional mutations that are described more particularly in Figure 5. Up to at least seven mutations are possible in a single mutein.

[0148] Other mutated nucleic acid sequences and proteins encoded by the mutated nucleic acid sequences are within the scope of the invention and are described more particularly in Example 5. These include the following: V482I/A484I (a double mutation); V366S/V482I/A484I (a triple mutation); G280T/G281S/V366S/V482I/A484I (a quintuple mutation); I294V/F296V/V366S/V482I/A484I (a quintuple mutation); I294V/F296V/V366S/V482I/A484I (a septuple mutation);

G280T/G281S/V366S/V480S/V482I (a quintuple mutation); I294V/F296V/366S/V480S/V482I (a quintuple mutation; I294V/F296V/V366S/V480S/A484I (a septuple mutation); S308T/V366S/V480S/V482I/A484I; S308T/V366S/V480S/A484I; S308T/V366S/V482I/A484I; S308T/V366S/V480S/V482I/; S308T/V366S/A484I;S308T/V366S/V482I; S308T/V366S/V480S; S308T/V366S; V366S/V480S/V482I/A484I; V366S/V480S/A484I; V366S/V482I/A484I (designated the M3 mutant); V366S/V480S/V482I; V366S/A484I; V366S/V480S; V480S/V482I/A484I; V482I/A484I; V480S/A484I; V480S/V482I;. L52E/M3 (where M3 is the triple mutant described above and the other mutations are included in an M3 background); L52E/G209E/M3; G209E/M3; C119S/M3; R113Q/M3; V109E/M3; E107D/C119S/M3; E107D/M3; P106M/C119S/M3; P106M/M3; L104V/C119S/M3; L104V/P106M/E107D/M3; L104V/M3; L103I/L104V/C119S/M3; L103I/L104V/M3; L103I/C119S/M3; L103I/M3; G280T/G281S/I294V/F296V/M3, I294V/F296V/M3, and G280T/G281S/M3. Other combinations of these mutations are within the scope of the invention.

[0149] Another aspect of the present invention relates to the generation of functional hydroxylases by domain swapping between the epiaristolochene hydroxylase (EAH) and HPO enzymes. These enzymes have a high degree of sequence identity and sequence similarity. Specifically, a total of six domains that are responsible for substrate specificity and/or recognition have been identified in the HPO enzyme (Example 5). These domains are designated SRS (for substrate recognition sequence), and are referred to herein as SRS 1, SRS 2, SRS 3, SRS 4, SRS 5, and SRS 6. Comparable domains are located in EAH. Accordingly, within the scope of the invention are functional hydroxylases generated by domain swapping that have either EAH activity (the production of capsidiol or 1β(OH)EA)) or HPO activity, or a combination of both activities. These include, but are not limited to: (1) a functional hydroxylase generated by domain swapping that includes SRS 1, SRS 2, and SRS 3 from HPO and SRS 4, SRS 5, and SRS 6 from EAH; (2) a functional hydroxylase generated by domain

swapping that includes SRS 1, SRS 2, SRS 3, and SRS 6 from HPO and SRS 4 and SRS 5 from EAH; a functional hydroxylase generated by domain swapping that includes SRS 1, SRS 2, SRS 3, SRS 4, and SRS 5 from EAH and SRS 6 from HPO; and a functional hydroxylase generated by domain swapping that includes SRS 1, SRS 2, SRS 3, and SRS 6 from EAH and SRS 4 and SRS 5 from HPO. All proteins according to the present invention produced by domain swapping include at least one SRS from HPO. Nucleic acids encoding these hydroxylases produced by domain swapping are also within the scope of the invention.

[0150] Hydroxylases (oxidases) produced by domain swapping and within the scope of the invention can be further modified by mutagenesis, particularly mutagenesis that results in the introduction of one or more point mutations such as are described above and set forth in greater detail in Figure 5. These muteins of hydrolases produced by domain swapping and subsequent mutagenesis are within the scope of the invention. Specific mutations are those of Example 5.

[0151] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0152] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is SEQ ID NO: 1.

[0153] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional

Hyoscyamus muticus premnaspirodiene oxidase protein, wherein the sequence encodes a protein of the sequence SEQ ID NO: 2.

[0154] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 97.5% identical to SEQ ID NO: 1.

[0155] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 99% identical to SEQ ID NO: 1.

[0156] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0157] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional

Hyoscyamus muticus premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to ten conservative amino acid substitutions from SEQ ID NO: 2.

[0158] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0159] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0160] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0161] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

[0162] In another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

[0163] In yet another embodiment, the present invention relates to a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

[0164] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0165] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is SEQ ID NO: 1.

[0166] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence encodes a protein of the sequence SEQ ID NO: 2.

[0167] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA

sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 97.5% identical to SEQ ID NO: 1.

[0168] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is at least 99% identical to SEQ ID NO: 1.

[0169] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0170] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to ten conservative amino acid substitutions from SEQ ID NO: 2.

[0171] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

[0172] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein, wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

[0173] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0174] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

[0175] In yet another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

[0176] In another embodiment, the present invention relates to a host cell transformed or transfected with a vector comprising an isolated and purified DNA sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

[0177] In another embodiment, the present invention relates to a method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

[0178] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

[0179] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein has the amino acid sequence of SEQ ID NO: 2.

[0180] In yet another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus*

premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0181] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0182] In another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity

of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 5 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0183] In yet another embodiment, the present invention relates to an isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone, wherein the protein ahs an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0184] In another embodiment, the present invention relates to an isolated and purified protein molecule having a functional terpene hydroxylase activity wherein the protein molecule has the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

[0185] In another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone.

[0186] In yet another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule of SEQ ID NO: 2 with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone.

[0187] In another embodiment, the present invention relates to a method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone, wherein the isolated and purified protein molecule has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

[0188] In another embodiment, the present invention relates to a yeast cell transformed or transfected with:

(a) a first vector including therein a DNA molecule encoding functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone;

(b) a second vector including therein a DNA molecule encoding a functional P450 reductase protein; and

(c) a third vector including therein a DNA molecule encoding a functional *Hyoscyamus muticus* premnaspirodiene protein synthase;

such that the yeast cell expresses: (1) the functional *Hyoscyamus muticus* premnaspirodiene oxidase protein in a quantity sufficient to hydroxylate valencene; (2) the functional P450 reductase protein in a quantity sufficient to supply reducing equivalents for the *Hyoscyamus muticus* premnaspirodiene oxidase protein; and (3) the functional *Hyoscyamus muticus* premnaspirodiene protein synthase in a quantity sufficient to produce premnaspirodiene; and such that the premnaspirodiene is converted by the cell to solavetivone.

[0189] In yet another embodiment, the present invention relates to a method of producing an oxidized terpene from unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to a host cell transformed or transfected with a vector comprising isolated and purified DNA sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0190] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence of SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0191] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA encoding a protein of SEQ ID NO: 2;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0192] In yet another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the DNA sequence is at least 95% identical to SEQ ID NO: 1 provided that the DNA sequence is translated into a protein encoding a functional *Hyoscyamus muticus* premnaspirodiene protein synthase protein such that the protein has a catalytic activity of successfully hydroxylating valencene at C2 first to nootkatol and then to nootkatone;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0193] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the sequence is at least 97.5% identical to SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0194] In another embodiment, the present invention relates to a method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell transformed or transfected with a vector comprising isolated and purified DNA sequence wherein the sequence is at least 99% identical to SEQ ID NO: 1;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

[0195] In yet another embodiment, the present invention relates to a method for producing a mutein of *Hyoscyamus muticus* premnaspirodiene oxidase with at least one altered property selected from the group consisting of regiospecificity and stereospecificity comprising the steps of:

(a) threading the wild-type *Hyoscyamus muticus* premnaspirodiene oxidase amino acid sequence onto the three-dimensional structure of a mammalian analogue enzyme;

(b) docking at least one sesquiterpene substrate into the predicted active site pocket;

(c) mapping amino acid residues within a defined distance of the modeled substrate molecules and rationalizing the amino acid residues with predicted chemical transformations catalyzed by the *Hyoscyamus muticus* premnaspirodiene oxidase to identify amino acid residues capable of steric, ionic, electronic and hydrophobic interactions with the substrate molecules;

(d) selecting at least one altered amino acid residue for its potential for substrate binding and positioning of the substrate molecule relative to the heme-catalytic center; and

(e) introducing the at least one altered amino acid amino acid residue by site-directed mutagenesis to produce the mutein.

[0196] Successive hydroxylation of valencene at C2, for example, yields nootkatone, a high demand, high value flavorant added to many of the commercial soft drinks sold worldwide. As discussed above, currently the practice of extracting nootkatone and other hydroxylated terpenes from plant or other biological materials, or organic synthesis of these compounds are considered expensive and unreliable processes.

[0197] The HPO gene now provides an alternative means of generating important stereochemically pure starting materials for the reliable and cost effective production of nootkatone and other high value sesquiterpenes. For example, the co-expression of a valencene synthase gene along with the HPO gene in transgenic plants or microbial cells could providing for large scale production of nootkatone. Alternatively, other terpene synthase genes could be used in combination with the HPO gene to generate other novel terpene moieties, which could be of value for pharmaceutical, agricultural and other industrial applications.

[0198] We are also claiming the utility of selective, site-directed mutagenesis of the HPO gene as a means for generating new biosynthetic .

capabilities in the resulting mutant enzyme. For example, mutation of the amino acid positions corresponding to valine 366, valine 482 and alanine 484 to serine, isoleucine and isoleucine, respectively, results in an enzyme that catalyzes hydroxylation at different regio-positions of the sesquiterpene substrates. We are furthermore claiming that coupled or co-expression of mutated sesquiterpene synthases in combination with mutated hydroxylase genes in heterologous hosts such as microbes or plants can generate additional novel chemical entities.

[0199] Figure 1 shows biosynthetic transformations catalyzed by the *Hyoscyamus muticus* premnaspirodiene synthase. Evidence provided below demonstrates that the HPO enzyme catalyzes the successive hydroxylation of sesquiterpene scaffolds, thus generating a first mono-hydroxylated form, followed by the subsequent ketone form. The HPO enzyme can utilize a broad range of structurally diverse substrates, such as the vetispirane (A) or eremophilane (B) class of sesquiterpenes.

[0200] The invention is described by the following Examples. These Examples are for illustrative purposes only and are not intended to limit the invention.

Example 1: Isolation of the HPO gene

[0201] The HPO gene was isolated using an RT-PCR strategy relying on the design of PCR primers derived from comparison to already cloned terpene hydroxylase genes, including the epi-aristolochene dihydroxylase gene (Ralston et al., 2001) and limonene-6-hydroxylase (Lupien et al., 1999). First stand cDNA was generated by reverse transcription of mRNA isolated from elicitor-induced *Hyoscyamus muticus* roots using oligo-dT as the initial primer, followed by PCR amplification using Forward primer 1 and Reverse primer 1 shown in Figure 2. Both RT and PCR reactions were performed under standard assay conditions. Subsequent studies using additional PCR primers identified below have confirmed the reproducible isolation of the HPO gene.

[0202] In Figure 2, PCR primers used to isolate the HPO gene were initially designed relative to the EAH gene (71D20, GenBank accession number AF368376), and subsequently to other 71D family members whose expression are inducible by biological and abiotic stress (71D4, accession # AJ296346; 71D6, U48434; 71D7, U48435; 71D16, AF166332) (SEQ ID NOs: 3-7 and 11-15). These GenBank accession numbers and the documentation associated with these numbers are incorporated herein by this reference. The starting position is indicated in Figure 2. Forward and reverse primers were F1 through F3 and R1 through R3. F1 is ATGCAATTCTTCAGCTTGGTTTCC (SEQ ID NO: 8). F2 is TTGGYTTCCATYTTCCTWTT (SEQ ID NO: 9). F3 is TTTYTGTTRAGGAAATGGAA (SEQ ID NO: 10). R1 is GGAATAGTTGGAAGAGCTCTCATT (SEQ ID NO: 16) R2 is TGAGGAATASTTGGAAGA (SEQ ID NO: 17). R3 is ACGGTGAGGAATASTTGGAA (SEQ ID NO: 18). In these sequences, Y is cytosine (C) or thymine (T). W is adenine (A) or thymine (T). S is guanine (G) or cytosine (C). R is adenine (A) or guanine (G).

[0203] The RT-PCR amplified DNA was cloned into the pGEM plasmid vector and sequenced according to standard procedures. The DNA sequence and deduced amino acid sequence for the encoded protein are shown below.

DNA sequence for the isolated HPO gene:

ATGCAATTCTTCAGCTTGGTTTCCATCTTCCTTTTTCTATCTTTTTTGTTTTTG
TTAAGGAAATGGAAGAACTCCAATAGCCAGTCCAAGAAATTGCCTCCAGGT
CCATGGAAACTTCCATTACTAGGTAGCATGCTTCATATGGTTGGTGGACTTC
CACATCATGTACTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCATCTT
CACTTGGTGAAGTTTCTGCTGTTGTTGTTACTTCTCCTGATATGGCAAAAGA
AGTACTAAAAACTCATGACATTGCGTTCGCGTCTAGGCCTAAACTTTTAGCC
CCAGAGATTGTATGTTACAACAGGTCTGACATTGCGTTTTGCCCTTATGGTG
ATTACTGGAGACAAATGCGTAAAATTTGTGTCTTGGAAGTGTTGAGTGCCAA
GAATGTTAGGTCATTCAGCTCTATTAGGCGCGATGAAGTGCTTCGTCTAGTT
AATTTTGTCCGATCATCTACGAGTGAGCCGGTTAACTTTACTGAAAGGCTGT

TTTTATTCACAAGTTCCATGACATGTAGATCAGCATTTGGGAAAGTGTTCAA
GGAACAGGAAACATTTATACAACTAATCAAAGAAGTGATAGGTTTAGCAGGA
GGATTTGATGTGGCTGACATCTTCCCATCACTGAAGTTCCTCCATGTACTAA
CTGGAATGGAGGGTAAGATTATGAAGGCTCACCATAAAGTAGATGCAATTGT
TGAGGATGTCATCAATGAGCATAAGAAGAACCTTGCAATGGGGAAAACTAAT
GGTGCATTAGGAGGTGAAGATCTAATTGATGTTCTTTTAAGACTTATGAATG
ATGGAGGCCTTCAATTTCCTATCACCAATGACAACATCAAAGCTATTATCTTT
GACATGTTTGCTGCTGGGACAGAGACTTCATCGTCAACACTTGTATGGGCTA
TGGTGCAAATGATGAGAAATCCAACTATACTAGCCAAAGCTCAAGCAGAAGT
AAGAGAAGCATTCAAAGGAAAAGAAACTTTCGATGAAAATGATGTCGAAGAG
TTGAAATACTTGAAGTTAGTCATTAAAGAAACTCTAAGACTCCATCCACCAGT
TCCACTTTTGGTCCCAAGAGAATGTAGGGAAGAAACAGAAATAAATGGCTAC
ACTATTCCAGTAAAGACCAAAGTCATGGTTAATGTTTGGGCATTAGGAAGAG
ATCCGAAATACTGGGACGACGCAGATAACTTCAAGCCAGAGAGATTTGAGC
AGTGTTCTGTGGACTTTATAGGTAACAATTTTGAATATCTTCCATTTGGTGGT
GGAAGGAGGATATGTCCAGGGATATCATTTGGTTTAGCTAATGTTTATTTGC
CATTGGCTCAATTGCTATATCATTTTGATTGGAAACTCCCTACTGGAATGGAA
CCAAAAGACTTGGATTTGACAGAATTGGTTGGAGTAACTGCTGCCAGAAAG
AGTGATCTTATGTTGGTTGCGACTCCTTATCAACCTTCTCGAGAGTAA (SEQ
ID NO: 1).

[0204] The amino acid sequence for the HPO protein deduced from the open reading frame of the HPO DNA sequence is:

MQFFSLVSIFLFLSFLFLLRKWKNSNSQSKKLPPGPWKLPLLGSMLHMVGGLPH
HVLRDLAKKYGPLMHLQLGEVSAVVVTSPDMAKEVLKTHDIAFASRPKLLAPEIV
CYNRSDIAFCPYGDYWRQMRKICVLEVLSAKNVRSFSSIRRDEVLRLVNFVRSS
TSEPVNFTERLFLFTSSMTCRSAFGKVFKEQETFIQLIKEVIGLAGGFDVADIFPS
LKFLHVLTGMEGKIMKAHHKVDAIVEDVINEHKKNLAMGKTNGALGGEDLIDVLL
RLMNDGGLQFPITNDNIKAIIFDMFAAGTETSSSTLVWAMVQMMRNPTILAKAQ
AEVREAFKGKETFDENDVEELKYLKLVIKETLRLHPPVPLLVPRECREETEINGY
TIPVKTKVMVNVWALGRDPKYWDDADNFKPERFEQCSVDFIGNNFEYLPFGG

GRRICPGISFGLANVYLPLAQLLYHFDWKLPTGMEPKDLDLTELVGVTAARKSD LMLVATPYQPSRE (SEQ ID NO: 2).

[0205] The HPO protein is highly homologous to previously characterized terpene hydroxylases. For example, HPO is 81% identical to EAH (Ralston et al., 2001) with approximately 91 amino acid substitutions, and 50% identical to limonene 6-hydroxylase (GenBank accession # AF124815, Lupien et al., 1999).

Example 2: Biochemical Characterization of the HPO gene

[0206] The HPO gene was functionally characterized by cloning the cDNA into a yeast expression vector (YePD-60), introducing the recombinant vector into yeast containing a suitable P450 reductase gene, such as line WAT11 (Urban et al., 1997), and inducing expression of the HPO gene by addition of galactose to the culture media (as per Ralston et al., 2001). The microsome fraction was subsequently isolated from the collected yeast cells and assayed for hydroxylase activity by incubation with sesquiterpene substrates and reducing equivalents (NADPH) for various lengths of time. The assays were then extracted with ethyl acetate and examined for substrate (sesquiterpene hydrocarbon) and hydroxylated reaction products by GC-MS. Figure 3 demonstrates the time dependent conversion of premnaspirodiene to solavetivol and solavetivone by microsomes from yeast over-expressing the HPO gene. The solavetivol and solavetivone reaction products were identified by identical retention time with authentic standards and by MS comparisons of the reaction products to authentic standards. Control assays, microsomes from yeast either not containing the HPO gene or yeast not induced to express the HPO gene, did not exhibit any conversion of premnaspirodiene to hydroxylated product.

[0207] In Figure 3, GC-MS analysis of reaction products generated by in vitro assays with microsomes from yeast over-expressing the HPO gene and incubated with premnaspirodiene as substrate is shown as a total ion chromatogram.

[0208] The HPO enzyme also exhibits broad substrate specificity, but

maintains regio-and stereo-specificity. For example, incubation of microsomes from yeast over-expressing the HPO gene with valencene yields time dependent biosynthesis of β-nootkatol and nootkatone (Fig. 4). Reaction products were identified by comparison of retention time and MS to authentic standards (Fig. 5).

[0209] In Figure 4.GC-MS analysis of reaction products generated by in vitro assays with microsomes from yeast over-expressing the HPO gene and incubated with valencene as substrate is shown as a total ion chromatogram.

[0210] In Figure 5, a comparison of mass spectra for reaction products generated in the in vitro assay (left panels) to authentic standards for ß-nootkatol and nootkatone (right panels) is shown (RT=retention time).

Example 3. Engineering Regio-and Stereo-Specificity of Terpene Hydroxylases

[0211] While the wildtype HPO enzyme exhibits broad substrate specificity, the regio-and stereo-specificity of the enzyme can be manipulated by selective site-directed mutagenesis. The HPO amino acid sequence was threaded onto the 3-Dimensional structure of the mammalian 2B4 hydroxylase (Scott et al., 2004) and several sesquiterpene substrate structures docked into the predicted active site pocket. Amino acid residues within 13 Å of the modeled substrate molecules were mapped and rationalized with predicted chemical transformations catalyzed by the HPO enzyme. Residues capable of steric, ionic, electronic and hydrophobic interactions with the substrate molecule were thus identified for their potential for substrate binding and positioning of the substrate molecule relative to the heme-catalytic center. These residues included V482, A484, V480 and others.

[0212] Site-directed mutagenesis was used to introduce changes into the HPO gene that resulted in amino acid substitutions within the HPO enzyme. For example, the codon corresponding to position 480 coding for valine was changed to that coding for serine, V480S. The mutant gene was over-expressed in yeast and isolated microsomes used in in vitro assays to assess the catalytic activity of

the mutant enzyme relative to the wildtype enzyme. Several examples of how site-directed mutagenesis was used to alter the catalytic activity and regiospecificity of the HPO enzyme are illustrated in Figure 6.

[0213] In Figure 6, site-directed mutagenesis was used to alter the DNA sequence of the HPO gene corresponding to those codons for the indicated amino acids, the mutant genes over-expressed in yeast and isolated microsomes used to biochemically characterize the mutant enzyme activities. The isolated microsomes were incubated with 5-epi-aristolochene as the substrate and the ethyl acetate extractable reaction products evaluated by GC-MS (Raltson et al., 2001).

[0214] IV. Coupling of HPO gene expression with that of terpene synthases in a heterologous host cell to generate a production platform for functionalized terpene compounds.

[0215] The HPO enzyme activity requires reducing equivalents in the form of electrons for catalytic activity, which are supplied by P450 reductase proteins. Yeast cells were first engineered with a P450 reductase gene, inserted into the genome of the yeast cell under the control of a strong expression promoter and using an auxotrophic selection marker such as the TRP1 gene (Urban et al., 1997). The HPO and *Hyoscyamus muticus* premnaspirodiene synthase, HPS, genes were then inserted into separate yeast expression vectors and transformed into yeast harboring the P450 reductase gene, creating yeast lines containing 3 engineered genes as illustrated in Figure 7A. The yeast lines were grown under standard culture conditions before addition of galactose to the culture media to induce expression of the engineered genes. Yeast cells were subsequently extracted with ethyl acetate and the organic extract examined by GC-MS. Figure 7B illustrates a typical GC chromatogram for the organic extracts and demonstrates the accumulation of premnaspirodiene, solavetivol and solavetivone by the engineered yeast cultures. Neither wildtype yeast nor engineered yeast lines grown without induction treatment accumulated any of

these compounds. Compounds were identified by MS comparisons to authentic standards.

[0216] Figure 7 is a cartoon depiction of the engineering of a yeast line to effect sesquiterpene hydrocarbon and oxygenated terpene biosynthesis (A). An ethyl acetate extract of P450 reductase, HPS and HPO engineered yeast cells assessed by GC-MS (B).

Additional experimentation has demonstrated the time-dependent accumulation of the sesquiterpene products. For example, premnaspirodiene tends to accumulate first and to the highest levels before either solavetivol or solavetivone accumulate. Accumulation occurs for over 210 hour of culture incubation.

Example 4: Ability of HPO to Utilize a Broad Range of Substrates

[0217] The assays of Example 4 were performed *in vitro* by overexpressing the HPO gene in a yeast host (WAT 11), isolating microsomes from the yeast, and then incubating microsomes with the indicated substrates.

[0218] Figure 8 shows the amino acid alignment between HPO (SEQ ID NO: 2) and EAH (SEQ ID NO: 19). The difference is 91 amino acids and two spaces.

[0219] Figure 9 shows a summary of the reactions catalyzed by HPO relative to EAH. The font size depicts the relative catalytic rates.

[0220] Figure 10 shows the CO difference spectrum for HPO expressed in yeast. To calculate exactly how much HPO protein is expressed and properly inserted into the membranes of yeast (the microsomal fraction), a CO (carbon monoxide) difference spectrum is obtained. The peak at 450 nm is specific for the heterologous expressed P450 gene and the absolute absorbance value can be used to calculate the absolute level of the HPO protein in the microsome preparation. This is necessary to calculate the kinetic constants of Figure 16,

below.

[0221] Figure 11 shows the reactions carried out with the substrate premnaspirodiene, indicating successive hydroxylation to 4β-solavetivol and then to solavetivone.

[0222] Figure 12 shows the reactions carried out with the substrate valencene, indicating successive hydroxylation to β-nootkatol and then to nootkatone. There is a side reaction resulting in the formation of α-nootkatol, but the relative specificity constant based on k_{cat}/K_m values favors the formation of β-nootkatol over α-nootkatol by a ratio of 70:1.

[0223] Figure 13 shows the reaction carried out with valencene with a reaction time of 1 minute.

[0224] Figure 14 shows the reaction carried out with 5-epi-aristolochene with a reaction time of 0, 1, 2, 5 or 10 minutes. The principal product is 2β(OH)EA, but there is a minor side product of 2α (OH)EA.

[0225] Figure 15 shows the reaction carried out with 4-epieremophilene and its double-bond isomer with a reaction time of 5 minutes (upper panel), as well as the reaction with α-cedrene with a reaction time of 5 minutes (lower panel).

[0226] Figure 16 is a table showing a comparison of enzyme kinetics of HPO for various substrates relative to the previously characterized EAH (5-epi-aristolochene dihydroxylase) hydroxylases. K_m, k_{cat}, and k_{cat}/K_m are shown.

Example 5: Generation of additional muteins having altered enzymatic activity

[0227] Our experimental premise is that if we can understand how an enzyme works and if we can define the role of specific amino acids in a catalytic cascade, then we can manipulate that enzyme in a direct manner for the production of chemical products of our own design.

[0228] The data in this example shows that we have successfully defined the positional information (which is broader than simply designating a particular amino acid residue at position x) that is important for both region- and stereo-specific hydroxylation in terpenes.

[0229] The strategy is to obtain stepwise mutants that gain new functionality.

[0230] A schematic depiction of a method for converting HPO into an enzyme with the same activity as EAH is shown in Figure 17.

[0231] Figure 18 shows the EAH activity of a number of muteins, including those produced by domain swapping between EAH and HPO, and those produced by domain swapping between EAH and HPO with additional mutations in the HPO domain, including: V482I/A484I (a double mutation); V366S/V482I/A484I (a triple mutation); G280T/G281S/V366S/V482I/A484I (a quintuple mutation); I294V/F296V/V366S/V482I/A484I (a quintuple mutation); I294V/F296V/V366S/V482I/A484I (a septuple mutation); G280T/G281S/V366S/V480S/V482I (a quintuple mutation); I294V/F296V/366S/V480S/V482I (a quintuple mutation; and I294V/F296V/V366S/V480S/A484I (a septuple mutation). Light bars indicate the EAH framework and dark bars indicate the HPO framework.

[0232] Figure 19 shows the general strategy of using domain-swapping mutations based on substrate recognition sequences (SRS), reciprocal site-directed mutagenesis based on homology modeling with mammalian P450s, and a combination of domain-swapping and site-directed mutagenesis. The six substrate recognition sites are shown along with EAH activity.

[0233] Figure 20 shows the results of homology modeling and site-directed mutagenesis in generating muteins of HPO indicated by the amino acid in the native (wildtype) enzyme, the amino acid position, then the mutant amino acid. For example, V366S indicates that the valine in position 366 of the wildtype enzyme has been mutated to serine. Mutants affecting SRS 4, 5, and 6 include: S308T/V366S/V480S/V482I/A484I; S308T/V366S/V480S/A484I; S308T/V366S/V482I/A484I; S308T/V366S/V480S/V482I/; S308T/V366S/A484I;S308T/V366S/V482I; and S308T/V366S/V480S. Mutants affecting SRS 4 and 5 include S308T/V366S. Mutants affecting SRS 5 and 6 include: V366S/V480S/V482I/A484I; V366S/V480S/A484I; V366S/V482I/A484I (designated the M3 mutant); V366S/V480S/V482I; V366S/A484I; and V366S/V480S. Mutants affecting SRS 5 include: V366S. Mutants affecting SRS6 include: V480S/V482I/A484I; V482I/A484I; V480S/A484I; V480S/V482I; A484I; V482I; and V480S. The 5-epiaristolochene hydroxylase activity of these enzymes is shown.

[0234] Figure 21 is similar to Figure 20, but depicts changes in SRS 1 and/or SRS 2 based on M3 (Figure 20). Mutants include: L52E/M3, L52E/G209E/M3, G209E/M3, C119S/M3, R113Q/M3, V109E/M3, E107D/C119S/M3, E107D/M3, P106M/C119S/M3, P106M/M3, L104V/C119S/M3, L104V/P106M/E107D/M3, L104V/M3, L103I/L104V/C119S/M3, L103I/L104V/M3, L103I/C119S/M3, and L103I/M3. The 5-epiaristolochene hydroxylase activity of these enzymes is shown.

[0235] Figure 22 shows the results from a combination of domain-

swapping mutations and site-directed mutagenesis; the mutation of V366S greatly diminishes the 5-epiaristolochene hydroxylase activity.

[0236] Figure 23 shows additional results from site-directed mutagenesis in SRS 4 as well as domain swapping. Mutants include: G280T/G281S/I294V/F296V/M3, I294V/F296V/M3, G280T/G281S/M3, V482I/A484I; V366S/V482I/A484I, and I294V/F296V/V366S/A482I/A484I. EAH activity is shown in terms of both 5-epiaristolochene hydroxylase activity and 1β(OH)EA hydroxylase activity.

[0237] References:

1. Lupien S, Karp F, Wildung M, Croteau R (1999) Regiospecific cytochrome P450 limonene hydroxylases from mint (Mentha) species: cDNA isolation, characterization, and functional expression of (-)-4S-limonene-3-hydroxylase and (-)-4S-limonene-6-hydroxylase. Archives Of Biochemistry And Biophysics 368: 181-192

2. Ralston L, Kwon ST, Schoenbeck M, Ralston J, Schenk DJ, Coates RM, Chappell J (2001) Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1,3-dihydroxylase from tobacco (Nicotiana tabacum). Archives Of Biochemistry And Biophysics 393: 222235

3. Scott EE, White MA, He YA, Johnson EF, Stout CD, Halpert JR (2004) Structure of mammalian cytochrome P4502B4 complexed with 4-(4-chlorophenyl) imidazole at 1.9-angstrom resolution -Insight into the range of P450 conformations and the coordination of redox partner binding. Journal Of Biological Chemistry 279: 27294-27301

4. Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related Arabidopsis thaliana NADPH-Cytochrome P450 reductases with P450 CYP73A5. Journal Of Biological Chemistry 272: 19176-19186

5. Back, K. W. & Chappell, J. Identifying functional domains within terpene cyclases using a domain-swapping strategy. Proceedings Of The National Academy Of Sciences Of The United States Of America 93, 6841-6845 (1996).

6. Mathis, J. R. et al. Pre-steady-state study of recombinant sesquiterpene cyclases. Biochemistry 36, 8340-8348 (1997).

7. Pompon, D., Louerat, B., Bronine, A. & Urban, P. in Cytochrome P450, Pt B 51-64 (1996).

8. Ralston, L. et al. Cloning, heterologous expression, and functional characterization of 5-epi-aristolochene-1,3-dihydroxylase from tobacco (Nicotiana tabacum). Archives Of Biochemistry And Biophysics 393, 222-235 (2001).

9. Rising, K. A., Starks, C. M., Noel, J. P. & Chappell, J. Demonstration of germacrene A as an intermediate in 5-epi-aristolochene synthase catalysis. Journal Of The American Chemical Society 122, 1861-1866 (2000).

10. Starks, C. M., Back, K. W., Chappell, J. & Noel, J. P. Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science 277, 1815-1820 (1997).

11. Takahashi, S. et al. Kinetic and molecular analysis of 5-epiaristolochene 1,3-dihydroxylase, a cytochrome P450 enzyme catalyzing successive hydroxylations of sesquiterpenes. Journal Of Biological Chemistry 280, 3686-3696 (2005).

12. Urban, P., Mignotte, C., Kazmaier, M., Delorme, F. & Pompon, D.

Cloning, yeast expression, and characterization of the coupling of two distantly related Arabidopsis thaliana NADPH-Cytochrome P450 reductases with P450 CYP73A5. Journal Of Biological Chemistry 272, 19176-19186 (1997).

ADVANTAGES OF THE INVENTION AND INDUSTRIAL APPLICABILITY

[0238] Nucleic acids and polypeptides according to the invention provide an efficient, rapid, and cost-effective route for the production of large quantities of oxidized terpenes, many of which are of high commercial value, including the beverage flavoring nootkatone. They obviate the necessity of isolating enzymes from natural products while increasing the uniformity of the process and reducing the reliance of the process on biological starting material.

[0239] The conversion of valencene to nootkatone is an industrial process resulting in a commercial and valuable product. Thus, nucleic acids, polypeptides, vectors, host cells, and processes according to the invention have industrial applicability.

[0240] With respect to ranges of values, the invention encompasses each intervening value between the upper and lower limits of the range to at least a tenth of the lower limit's unit, unless the context clearly indicates otherwise. Moreover, the invention encompasses any other stated intervening values and ranges including either or both of the upper and lower limits of the range, unless specifically excluded from the stated range.

[0241] Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of ordinary skill in the art to which this invention belongs. One of ordinary skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test this invention.

[0242] The publications and patents discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

[0243] All the publications cited are incorporated herein by reference in their entireties, including all published patents, patent applications, literature references, as well as those publications that have been incorporated in those published documents. However, to the extent that any publication incorporated herein by reference refers to information to be published, applicants do not admit that any such information published after the filing date of this application to be prior art.

[0244] As used in this specification and in the appended claims, the singular forms include the plural forms. For example the terms "a," "an," and "the" include plural references unless the content clearly dictates otherwise. Additionally, the term "at least" preceding a series of elements is to be understood as referring to every element in the series. The inventions illustratively described herein can suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising," "including," "containing," etc. shall be read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the future shown and described or any portion thereof, and it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the inventions herein disclosed can be

resorted by those skilled in the art, and that such modifications and variations are considered to be within the scope of the inventions disclosed herein. The inventions have been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the scope of the generic disclosure also form part of these inventions. This includes the generic description of each invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised materials specifically resided therein. In addition, where features or aspects of an invention are described in terms of the Markush group, those schooled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group. It is also to be understood that the above description is intended to be illustrative and not restrictive. Many embodiments will be apparent to those of in the art upon reviewing the above description. The scope of the invention should therefore, be determined not with reference to the above description, but should instead be determined with reference to the appended claims, along with the full scope of equivalents to which such claims are entitled. Those skilled in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described. Such equivalents are intended to be encompassed by the following claims.

We claim:

1. An isolated and purified nucleic acid sequence encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

2. The isolated and purified nucleic acid sequence of claim 1 that is DNA.

3. The isolated and purified DNA sequence of claim 2 wherein the sequence is

ATGCAATTCTTCAGCTTGGTTTCCATCTTCCTTTTTCTATCTTTTTTGTTTTTG
TTAAGGAAATGGAAGAACTCCAATAGCCAGTCCAAGAAATTGCCTCCAGGT
CCATGGAAACTTCCATTACTAGGTAGCATGCTTCATATGGTTGGTGGACTTC
CACATCATGTACTTAGAGATTTAGCAAAAAAATATGGACCACTTATGCATCTT
CAACTTGGTGAAGTTTCTGCTGTTGTTGTTACTTCTCCTGATATGGCAAAAG
AAGTACTAAAAACTCATGACATTGCGTTCGCGTCTAGGCCTAAACTTTTAGC
CCCAGAGATTGTATGTTACAACAGGTCTGACATTGCGTTTTGCCCTTATGGT
GATTACTGGAGACAAATGCGTAAAATTTGTGTCTTGGAAGTGTTGAGTGCCA
AGAATGTTAGGTCATTCAGCTCTATTAGGCGCGATGAAGTGCTTCGTCTAGT
TAATTTTGTCCGATCATCTACGAGTGAGCCGGTTAACTTTACTGAAAGGCTG
TTTTTATTCACAAGTTCCATGACATGTAGATCAGCATTTGGGAAAGTGTTCAA
GGAACAGGAAACATTTATACAACTAATCAAAGAAGTGATAGGTTTAGCAGGA
GGATTTGATGTGGCTGACATCTTCCCATCACTGAAGTTCCTCCATGTACTAA
CTGGAATGGAGGGTAAGATTATGAAGGCTCACCATAAAGTAGATGCAATTGT
TGAGGATGTCATCAATGAGCATAAGAAGAACCTTGCAATGGGGAAAACTAAT
GGTGCATTAGGAGGTGAAGATCTAATTGATGTTCTTTTAAGACTTATGAATG
ATGGAGGCCTTCAATTTCCTATCACCAATGACAACATCAAAGCTATTATCTTT
GACATGTTTGCTGCTGGGACAGAGACTTCATCGTCAACACTTGTATGGGCTA
TGGTGCAAATGATGAGAAATCCAACTATACTAGCCAAAGCTCAAGCAGAAGT

AAGAGAAGCATTCAAAGGAAAAGAAACTTTCGATGAAAATGATGTCGAAGAG
TTGAAATACTTGAAGTTAGTCATTAAAGAAACTCTAAGACTCCATCCACCAGT
TCCACTTTTGGTCCCAAGAGAATGTAGGGAAGAAACAGAAATAAATGGCTAC
ACTATTCCAGTAAAGACCAAAGTCATGGTTAATGTTTGGGCATTAGGAAGAG
ATCCGAAATACTGGGACGACGCAGATAACTTCAAGCCAGAGAGATTTGAGC
AGTGTTCTGTGGACTTTATAGGTAACAATTTTGAATATCTTCCATTTGGTGGT
GGAAGGAGGATATGTCCAGGGATATCATTTGGTTTAGCTAATGTTTATTTGC
CATTGGCTCAATTGCTATATCATTTTGATTGGAAACTCCCTACTGGAATGGAA
CCAAAAGACTTGGATTTGACAGAATTGGTTGGAGTAACTGCTGCCAGAAAG
AGTGATCTTATGTTGGTTGCGACTCCTTATCAACCTTCTCGAGAGTAA (SEQ
ID NO: 1).

4. The isolated and purified DNA sequence of claim 2 wherein the sequence encodes a protein of the sequence
MQFFSLVSIFLFLSFLFLLRKWKNSNSQSKKLPPGPWKLPLLGSMLHMVGGLPH
HVLRDLAKKYGPLMHLQLGEVSAVVVTSPDMAKEVLKTHDIAFASRPKLLAPEIV
CYNRSDIAFCPYGDYWRQMRKICVLEVLSAKNVRSFSSIRRDEVLRLVNFVRSS
TSEPVNFTERLFLFTSSMTCRSAFGKVFKEQETFIQLIKEVIGLAGGFDVADIFPS
LKFLHVLTGMEGKIMKAHHKVDAIVEDVINEHKKNLAMGKTNGALGGEDLIDVLL
RLMNDGGLQFPITNDNIKAIIFDMFAAGTETSSSTLVWAMVQMMRNPTILAKAQ
AEVREAFKGKETFDENDVEELKYLKLVIKETLRLHPPVPLLVPRECREETEINGY
TIPVKTKVMVNVWALGRDPKYWDDADNFKPERFEQCSVDFIGNNFEYLPFGG
GRRICPGISFGLANVYLPLAQLLYHFDWKLPTGMEPKDLDLTELVGVTAARKSD
LMLVATPYQPSRE (SEQ ID NO: 2).

5. The isolated and purified nucleic acid sequence of claim 1 wherein the sequence is a nucleic acid sequence that is at least 95% identical to SEQ ID NO: 1 provided that the nucleic acid sequence is translated into a protein encoding a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

6. The isolated and purified nucleic acid sequence of claim 5 that is DNA.

7. The isolated and purified nucleic acid sequence of claim 5 wherein the sequence is a nucleic acid sequence that is at least 97.5% identical to SEQ ID NO: 1.

8. The isolated and purified nucleic acid sequence of claim 7 that is DNA.

9. The isolated and purified nucleic acid sequence of claim 7 wherein the sequence is a nucleic acid sequence that is at least 99% identical to SEQ ID NO: 1.

10. The isolated and purified nucleic acid sequence of claim 9 that is DNA.

11. The isolated and purified DNA sequence of claim 2 wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2, wherein the protein encoded is a functional *Hyoscyamus muticus* premnaspirodiene synthase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone and wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

12. The isolated and purified DNA sequence of claim 11 wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2.

13. The isolated and purified DNA sequence of claim 12 wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to five conservative amino acid substitutions from SEQ ID NO: 2.

14. The isolated and purified DNA sequence of claim 13 wherein the sequence is selected from the group consisting of:

(a) SEQ ID NO: 1; and

(b) a DNA sequence encoding a protein differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2.

15. An isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

16. The isolated and purified nucleic acid sequence of claim 15 wherein the nucleic acid is DNA.

17. An isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 482 is changed from valine to isoleucine.

18. The isolated and purified nucleic acid sequence of claim 17 wherein the nucleic acid is DNA.

19. An isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 484 is changed from alanine to isoleucine.

20. The isolated and purified nucleic acid sequence of claim 19 wherein the nucleic acid is DNA.

21. An isolated and purified nucleic acid sequence encoding a protein having the sequence of SEQ ID NO: 2 wherein residue 366 is changed from valine to serine.

22. The isolated and purified nucleic acid sequence of claim 21 wherein the nucleic acid is DNA.

23. A vector comprising the isolated and purified DNA of claim 2.

24. A vector comprising the isolated and purified DNA of claim 3.

25. A vector comprising the isolated and purified DNA of claim 4.

26. A vector comprising the isolated and purified DNA of claim 6.

27. A vector comprising the isolated and purified DNA of claim 8.

28. A vector comprising the isolated and purified DNA of claim 10.

29. A vector comprising the isolated and purified DNA of claim 11.

30. A vector comprising the isolated and purified DNA of claim 14.

31. A vector comprising the isolated and purified DNA of claim 16.

32. A vector comprising the isolated and purified DNA of claim 18.

33. A vector comprising the isolated and purified DNA of claim 20.

34. A vector comprising the isolated and purified DNA of claim 22.

35. A host cell transformed or transfected with the vector of claim 23.

36. The host cell of claim 35 that is eukaryotic.

37. The host cell of claim 36 that is a yeast cell.

38. A host cell transformed or transfected with the vector of claim 24.

39. The host cell of claim 38 that is eukaryotic.

40. The host cell of claim 39 that is a yeast cell.

41. A host cell transformed or transfected with the vector of claim 25.

42. The host cell of claim 41 that is eukaryotic.

43. The host cell of claim 42 that is a yeast cell.

44. A host cell transformed or transfected with the vector of claim 26.

45. The host cell of claim 44 that is eukaryotic.

46. The host cell of claim 45 that is a yeast cell.

47. A host cell transformed or transfected with the vector of claim 27.

48. The host cell of claim 47 that is eukaryotic.

49. The host cell of claim 48 that is a yeast cell.

50. A host cell transformed or transfected with the vector of claim 28.

51. The host cell of claim 50 that is eukaryotic.

52. The host cell of claim 51 that is a yeast cell.

53. A host cell transformed or transfected with the vector of claim 29.

54. The host cell of claim 53 that is eukaryotic.

55. The host cell of claim 53 that is a yeast cell.

56. A host cell transformed or transfected with the vector of claim 30.

57. The host cell of claim 56 that is eukaryotic.

58. The host cell of claim 57 that is a yeast cell.

59. A host cell transformed or transfected with the vector of claim 31.

60. The host cell of claim 59 that is eukaryotic.

61. The host cell of claim 60 that is a yeast cell.

62. A host cell transformed or transfected with the vector of claim 32.

63. The host cell of claim 62 that is eukaryotic.

64. The host cell of claim 63 that is a yeast cell.

65. A host cell transformed or transfected with the vector of claim 33.

66. The host cell of claim 65 that is eukaryotic.

67. The host cell of claim 66 that is a yeast cell.

68. A host cell transformed or transfected with the vector of claim 34.

69. The host cell of claim 68 that is eukaryotic.

70. The host cell of claim 69 that is a yeast cell.

71. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 35 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

72. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 38 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

73. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein

has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 41 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

74. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 44 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

75. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 47 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

76. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein

has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 50 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

77. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 53 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

78. A method of producing an isolated protein having *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone comprising the steps of:

(a) culturing the host cell of claim 56 under conditions wherein a protein having *Hyoscyamus muticus* premnaspirodiene synthase activity is expressed by the host cell; and

(b) isolating the protein having *Hyoscyamus muticus* premnaspirodiene synthase activity so that isolated protein is produced.

79. A method of producing an isolated protein having terpene synthase activity comprising the steps of:

(a) culturing the host cell of claim 59 under conditions wherein a protein having terpene synthase activity is expressed by the host cell; and

(b) isolating the protein having terpene synthase activity so that isolated protein is produced.

80. A method of producing an isolated protein having terpene synthase activity comprising the steps of:

(a) culturing the host cell of claim 62 under conditions wherein a protein having terpene synthase activity is expressed by the host cell; and

(b) isolating the protein having terpene synthase activity so that isolated protein is produced.

81. A method of producing an isolated protein having terpene synthase activity comprising the steps of:

(a) culturing the host cell of claim 65 under conditions wherein a protein having terpene synthase activity is expressed by the host cell; and

(b) isolating the protein having terpene synthase activity so that isolated protein is produced.

82. A method of producing an isolated protein having terpene synthase activity comprising the steps of:

(a) culturing the host cell of claim 68 under conditions wherein a protein having terpene synthase activity is expressed by the host cell; and

(b) isolating the protein having terpene synthase activity so that isolated protein is produced.

83. An isolated and purified protein molecule having functional *Hyoscyamus muticus* premnaspirodiene synthase activity such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone.

84. The isolated and purified protein molecule of claim 83 wherein the protein has the amino acid sequence of SEQ ID NO: 2.

85. The isolated and purified protein molecule of claim 83 wherein the protein has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

86. The isolated and purified protein molecule of claim 85 wherein the protein has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 10 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

87. The isolated and purified protein molecule of claim 86 wherein the protein has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 5 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

88. The isolated and purified protein molecule of claim 87 wherein the protein has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to two conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

89. An isolated and purified protein molecule having a functional terpene hydroxylase activity wherein the protein molecule has the sequence of SEQ ID NO: 2 wherein residue 480 is changed from valine to serine.

90. A method of producing nootkatone comprising the step of reacting the isolated and purified protein molecule of claim 83 with valencene under conditions in which the protein molecule catalyzes the successive oxidation of the valencene first to nootkatol and then to nootkatone.

91. The method of claim 90 wherein the isolated and purified protein molecule has the amino acid sequence of SEQ ID NO: 2.

92. The method of claim 90 wherein the isolated and purified protein molecule has an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO: 2; and

(b) an amino acid sequence differing by from one to 20 conservative amino acid substitutions from SEQ ID NO: 2 wherein a conservative amino acid substitution is one of the following substitutions: Ala/Gly or Ser; Arg/Lys; Asn/Gln or His; Asp/Glu; Cys/Ser; Gln/Asn; Gly/Asp; Gly/Ala or Pro; His/Asn or Gln; Ile/Leu or Val; Leu/Ile or Val; Lys/Arg or Gln or Glu; Met/Leu or Tyr or Ile; Phe/Met or Leu or Tyr; Ser/Thr; Thr/Ser; Trp/Tyr; Tyr/Trp or Phe; Val/Ile or Leu.

93. A yeast cell transformed or transfected with:

(a) a first vector including therein a DNA molecule encoding a functional *Hyoscyamus muticus* premnaspirodiene oxidase protein such that the protein has a catalytic activity of successively hydroxylating valencene at C2 first to nootkatol and then to nootkatone;

(b) a second vector including therein a DNA molecule encoding a functional P450 reductase protein; and

(c) a third vector including therein a DNA molecule encoding a functional *Hyoscyamus muticus* premnaspirodiene protein synthase;

such that the yeast cell expresses: (1) the functional *Hyoscyamus muticus* premnaspirodiene oxidase protein in a quantity sufficient to hydroxylate valencene; (2) the functional P450 reductase protein in a quantity sufficient to supply reducing equivalents for the *Hyoscyamus muticus* premnaspirodiene oxidase protein; and (3) the functional *Hyoscyamus muticus* premnaspirodiene protein synthase in a quantity sufficient to produce premnaspirodiene; and such that the premnaspirodiene is converted by the cell to solavetivone.

94. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 35;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

95. The method of claim 94 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

96. The method of claim 94 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

97. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 38;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

98. The method of claim 97 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

99. The method of claim 97 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

100. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 41;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

101. The method of claim 100 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

102. The method of claim 100 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

103. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 44;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

104. The method of claim 103 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

105. The method of claim 103 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

106. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 47;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

107. The method of claim 106 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

108. The method of claim 106 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

109. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 50;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

110. The method of claim 109 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

111. The method of claim 109 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

112. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 53;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

113. The method of claim 112 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

114. The method of claim 112 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

115. A method of producing an oxidized terpene from an unoxidized terpene substrate comprising the steps of:

(a) providing the unoxidized terpene substrate to the host cell of claim 56;

(b) culturing the host cell under conditions in which the unoxidized terpene substrate is oxidized; and

(c) isolating the oxidized terpene.

116. The method of claim 115 wherein the unoxidized terpene substrate is valencene and the oxidized terpene is nootkatone.

117. The method of claim 115 wherein the unoxidized terpene substrate is premnaspirodiene and the oxidized terpene is solavetivone.

118. A method for producing a mutein of *Hyoscyamus muticus* premnaspirodiene oxidase with at least one altered property selected from the group consisting of regiospecificity and stereospecificity comprising the steps of:

(a) threading the wild-type *Hyoscyamus muticus* premnaspirodiene oxidase amino acid sequence onto the three-dimensional structure of a mammalian analogue enzyme;

(b) docking at least one sesquiterpene substrate into the predicted active site pocket;

(c) mapping amino acid residues within a defined distance of the modeled substrate molecules and rationalizing the amino acid residues with

predicted chemical transformations catalyzed by the *Hyoscyamus muticus* premnaspirodiene oxidase to identify amino acid residues capable of steric, ionic, electronic and hydrophobic interactions with the substrate molecules;

(d) selecting at least one altered amino acid residue for its potential for substrate binding and positioning of the substrate molecule relative to the heme-catalytic center; and

(e) introducing the at least one altered amino acid amino acid residue by site-directed mutagenesis to produce the mutein.

119. The method of claim 118 wherein the mammalian analogue enzyme is mammalian 2B4 hydroxylase.

120. The method of claim 118 wherein the defined distance is about 13 Å.

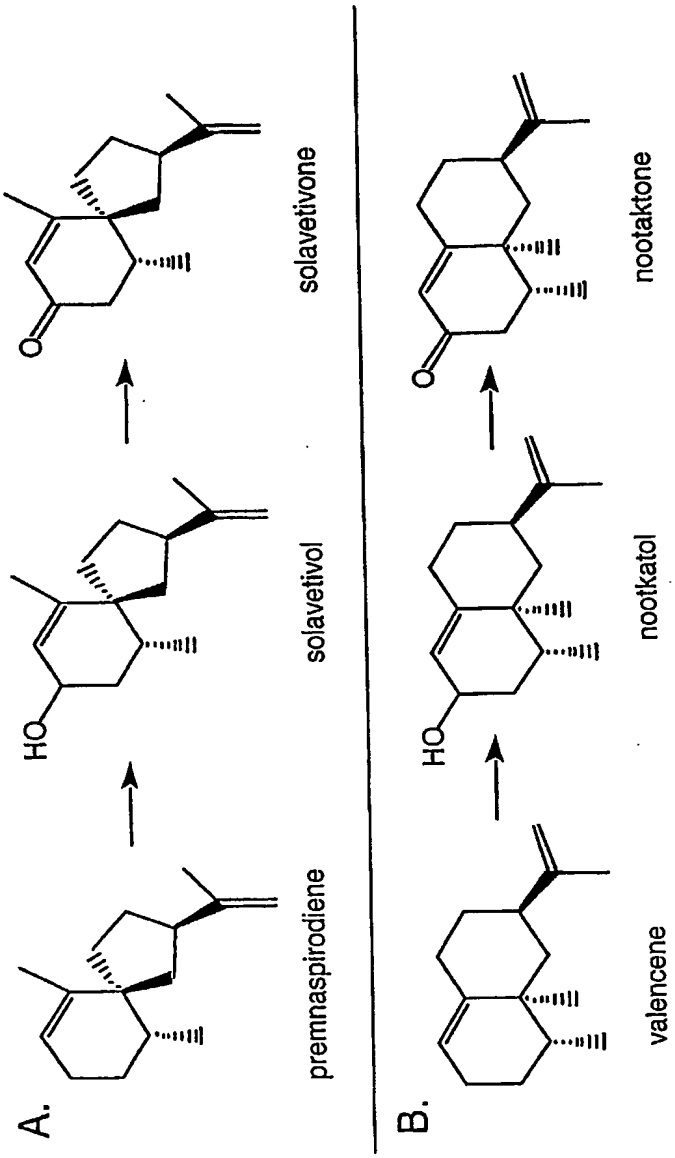

A.
premnaspirodiene
solavetivol
solavetivone
HO
O
B.
valencene
nootkatol
nootaktone
HO
O

Figure 1

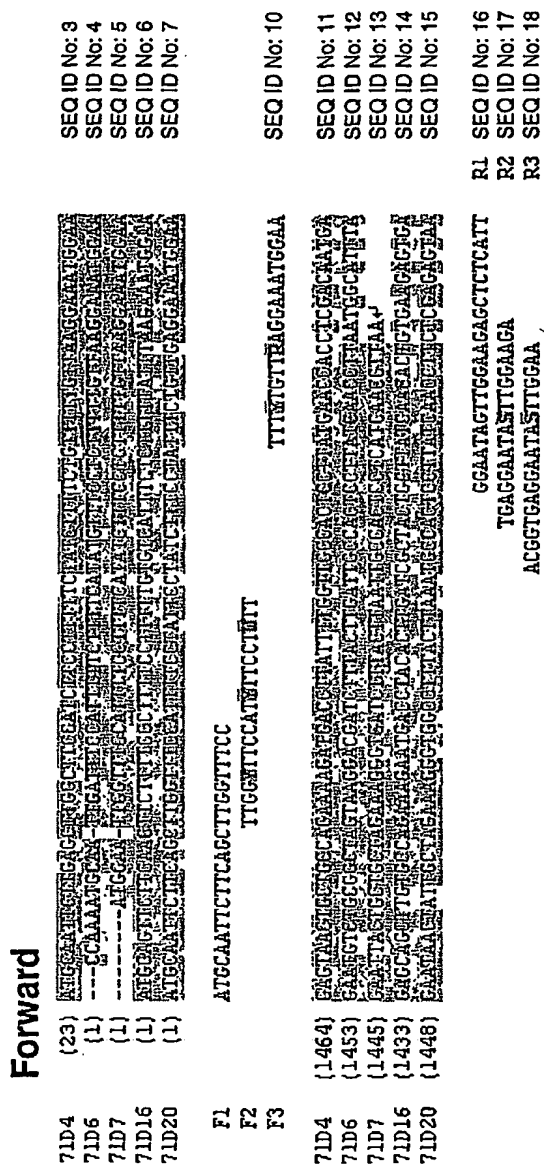

Forward
SEQ ID No: 3
SEQ ID No: 4
SEQ ID No: 5
SEQ ID No: 6
SEQ ID No: 7
SEQ ID No: 10
SEQ ID No: 11
SEQ ID No: 12
SEQ ID No: 13
SEQ ID No: 14
SEQ ID No: 15
R1 SEQ ID No: 16
R2 SEQ ID No: 17
R3 SEQ ID No: 18
Reverse

Figure 2

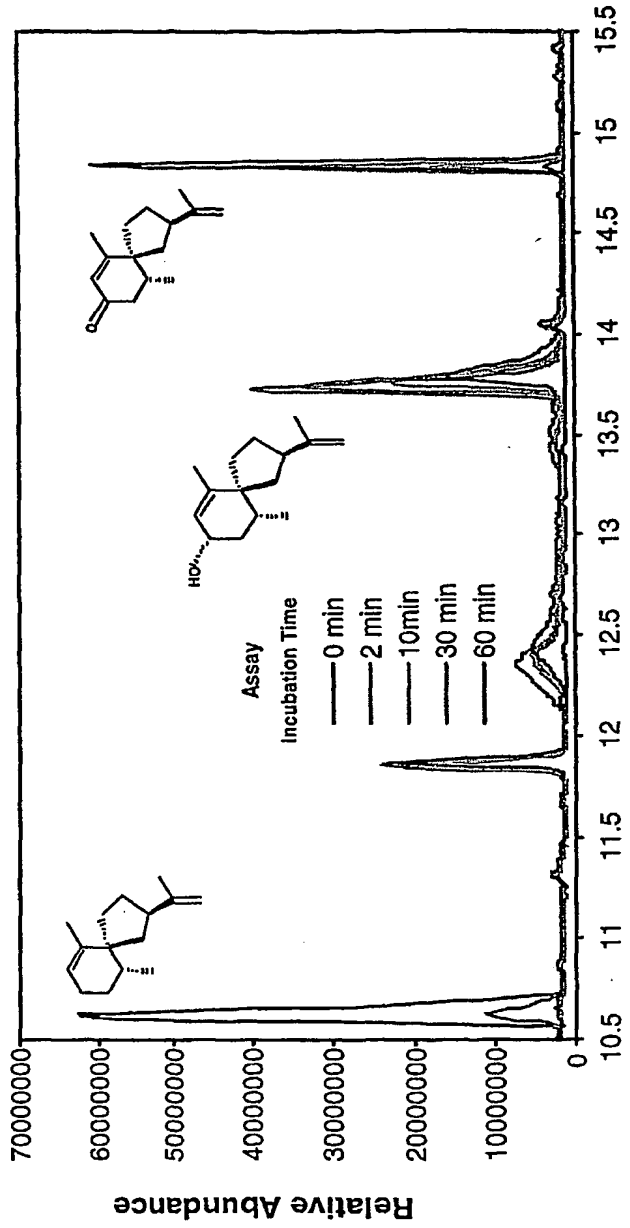

Relative Abundance
70000000
60000000
50000000
40000000
30000000
20000000
10000000
0
10.5
11
11.5
12
12.5
13
13.5
14
14.5
15
15.5
Assay
Incubation Time
0 min
2 min
10min
30 min
60 min
HO
O

Figure 3

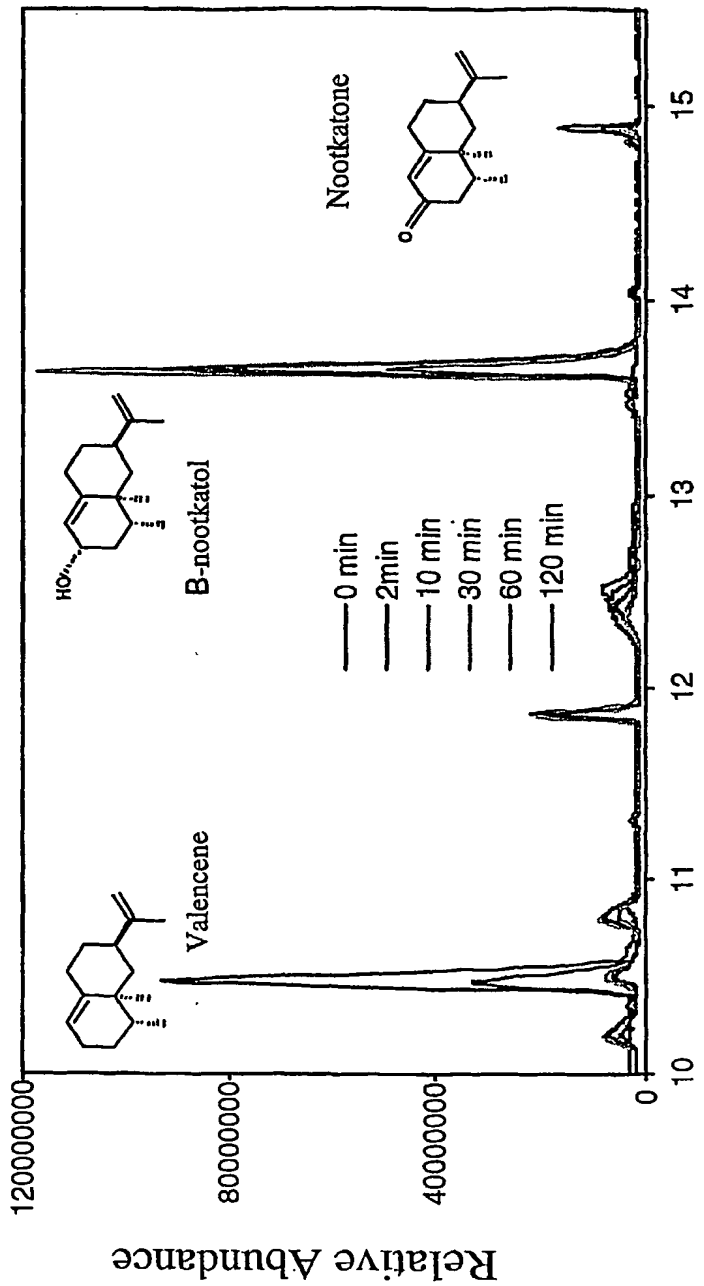

Relative Abundance
120000000
80000000
40000000
0
10
11
12
13
14
15
Retention time (min.)
Valencene
B-nootkatol
Nootkatone
0 min
2min
10 min
30 min
60 min
120 min

Figure 4

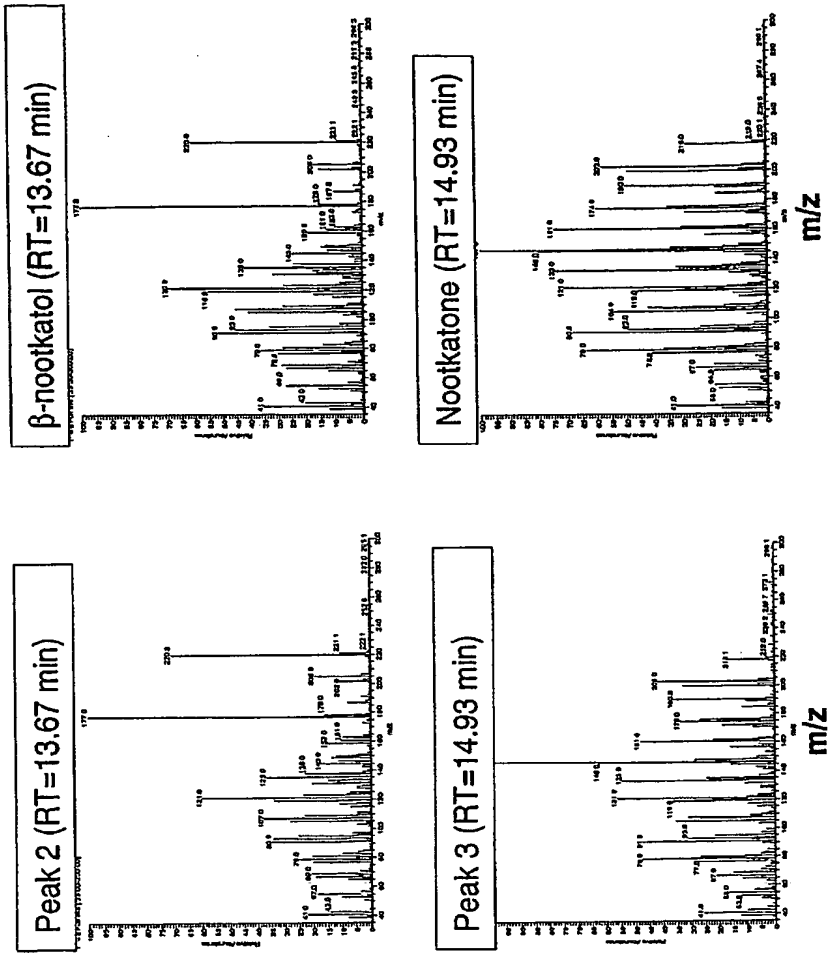

β-nootkatol (RT=13.67 min)
Nootkatone (RT=14.93 min)
Peak 2 (RT=13.67 min)
Peak 3 (RT=14.93 min)
m/z
m/z

Figure 5

5/24

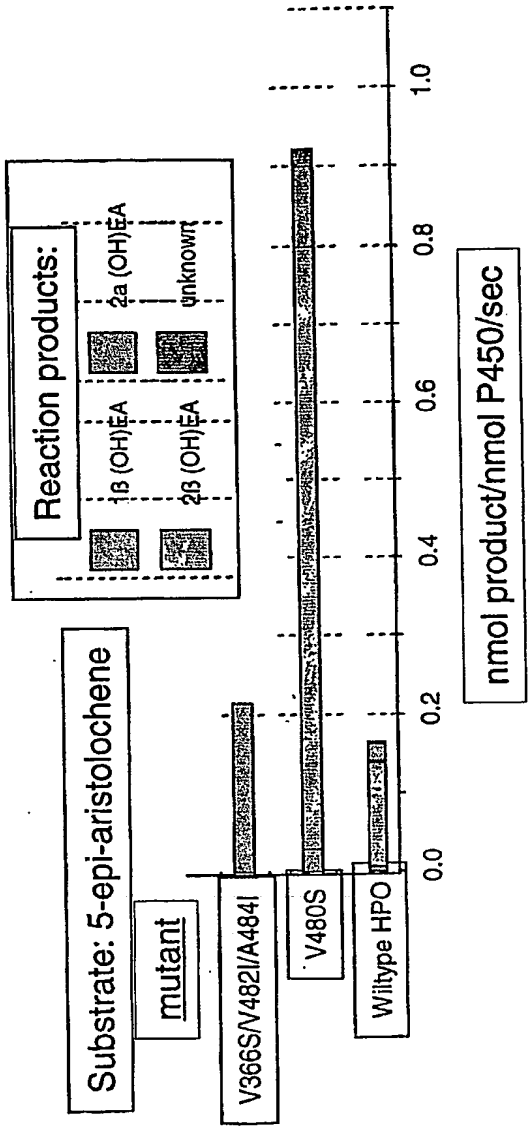

Substrate: 5-epi-aristolochene
mutant
V366S/V482I/A484I
V480S
Wiltype HPO
0.0
0.2
0.4
0.6
0.8
1.0
nmol product/nmol P450/sec
Reaction products:
1β (OH)EA
2β (OH)EA
2a (OH)EA
unknown

Figure 6

6/24

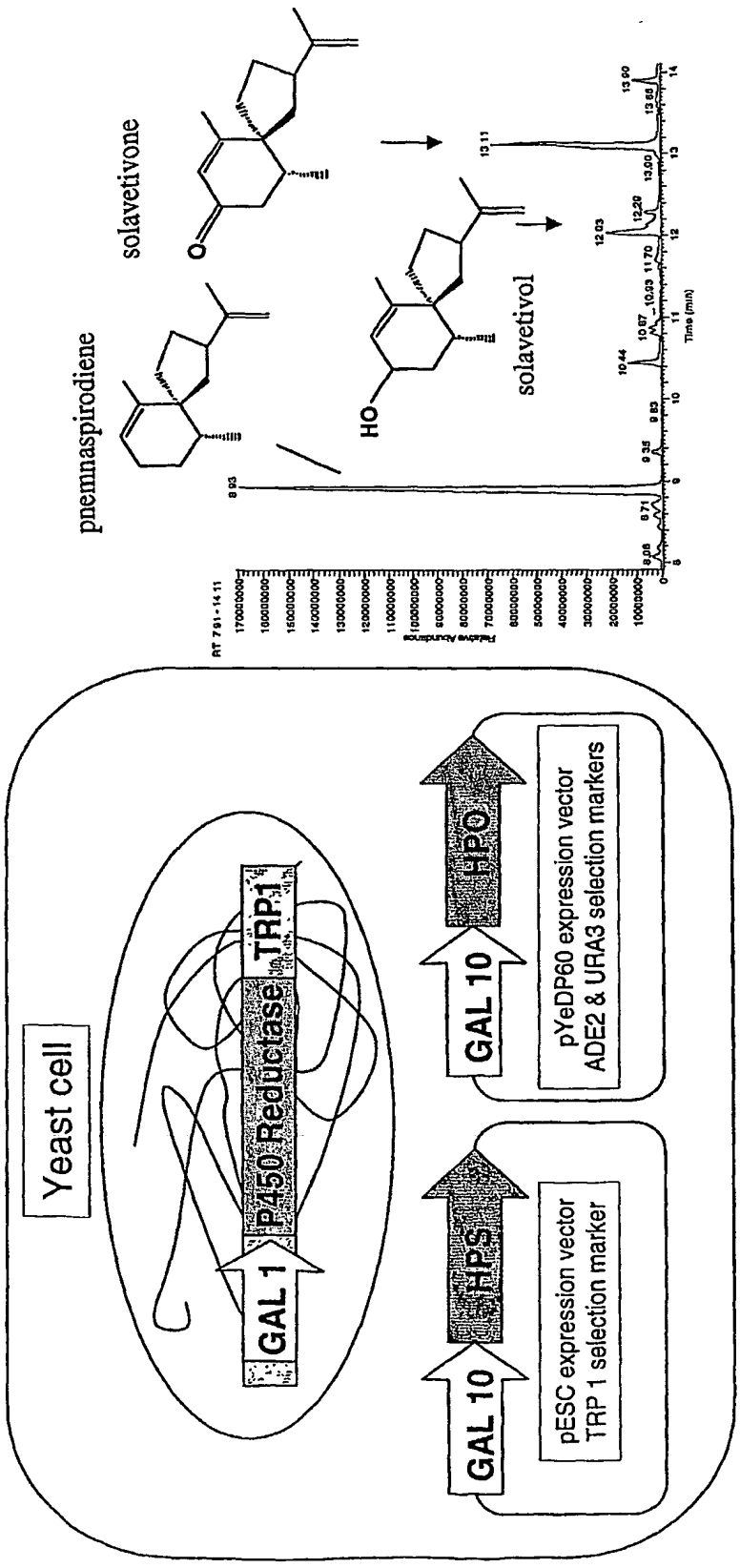

pnemnaspirodiene
solavetivone
solavetivol
HO
O
Yeast cell
GAL 1
P450 Reductase
TRP1
GAL 10
HPS
pESC expression vector
TRP 1 selection marker
GAL 10
HPO
pYeDP60 expression vector
ADE2 & URA3 selection markers

Figure 7

7/24

Amino acids alignment between EAH and HPO

```
HPH     MQFFSLVSIF LFLSFLFLLR KWKNSNSQSK KLPPGPWKLP LLGSMLHMVG
TEAH    MQFFSLVSIF LFLSFLFLLR KWKNSNSQSK KLPPGPWKIP ILGSMLHMIG

HPH     GLPHHVLRDL AKKYGPLMHL QLGEVSAVVV TSPDMAKEVL KTHDIAFASR
TEAH    GEPHHVLRDL AKKYGPLMHL QLGEISAVVV TSRDMAKEVL KTHDVVFASR

HPH     PKLLAPEIVC YNRSDIAFCP YGDYWRQMRK ICVLEVLSAK NVRSFSSIRR
TEAH    PKIVAMDIIC YNQSDIAFSP YGDHWRQMRK ICVMELLNAK NVRSFSSIRR

HPH     DEVLRLVNFV R..SSTSEPV NFTERLFLFT SSMTCRSAFG KVFKEQETFI
TEAH    DEVVRLIDSI RSDSSSGELV NFTQRIIWFA SSMTCRSAFG QVLKGQDIFA

HPH     QLIKEVIGLA GGFDVADIFP SLKFLHVLTG MEGKIMKAHH KVDAIVEDVI
TEAH    KKIREVIGLA EGFDVVDIFP TYKFLHVLSG MKRKLLNAHL KVDAIVEDVI

HPH     NEHKKNLAMG KTNGALGGED LIDVLLRLMN DGGLQFPITN DNIKAIIFDM
TEAH    NEHKKNLAAG KSNGALGGED LIDVLLRLMN DTSLQFPITN DNIKAVIVDM

HPH     FAAGTETSSS TLVWAMVQMM RNPTILAKAQ AEVREAFKGK ETFDENDVEE
TEAH    FAAGTETSST TTVWAMAEMM KNPSVFTKAQ AEVREAFRDK VSFDENDVEE

HPH     LKYLKLVIKE TLRLHPPVPL LVPRECREET EINGYTIPVK TKVMVNVWAL
TEAH    LKYLKLVIKE TLRLHPPSPL LVPRECREDT DINGYTIPAK TKVMVNVWAL

HPH     GRDPKYWDDA DNFKPERFEQ CSVDFIGNNF EYLPFGGGRR ICPGISFGLA
TEAH    GRDPKYWDDA ESFKPERFEQ CSVDFFGNNF EFLPFGGGRR ICPGMSFGLA

HPH     NVYLPLAQLL YHFDWKLPTG MEPKDLDLTE LVGVTAARKS DLMLVATPYQ
TEAH    NLYLPLAQLL YHFDWKLPTG IMPRDLDLTE LSGITIARKG GLYLNATPYQ

HPH     PSRE (SEQ ID No: 2)
TEAH    PSRE (SEQ ID No: 19)
```

Difference between HPO and EAH: 91 amino acids and 2 spaces

FIG. 8

Summary of the Reactions catalyzed by HPO relative to EAH
Font size depicts relative catalytic rates

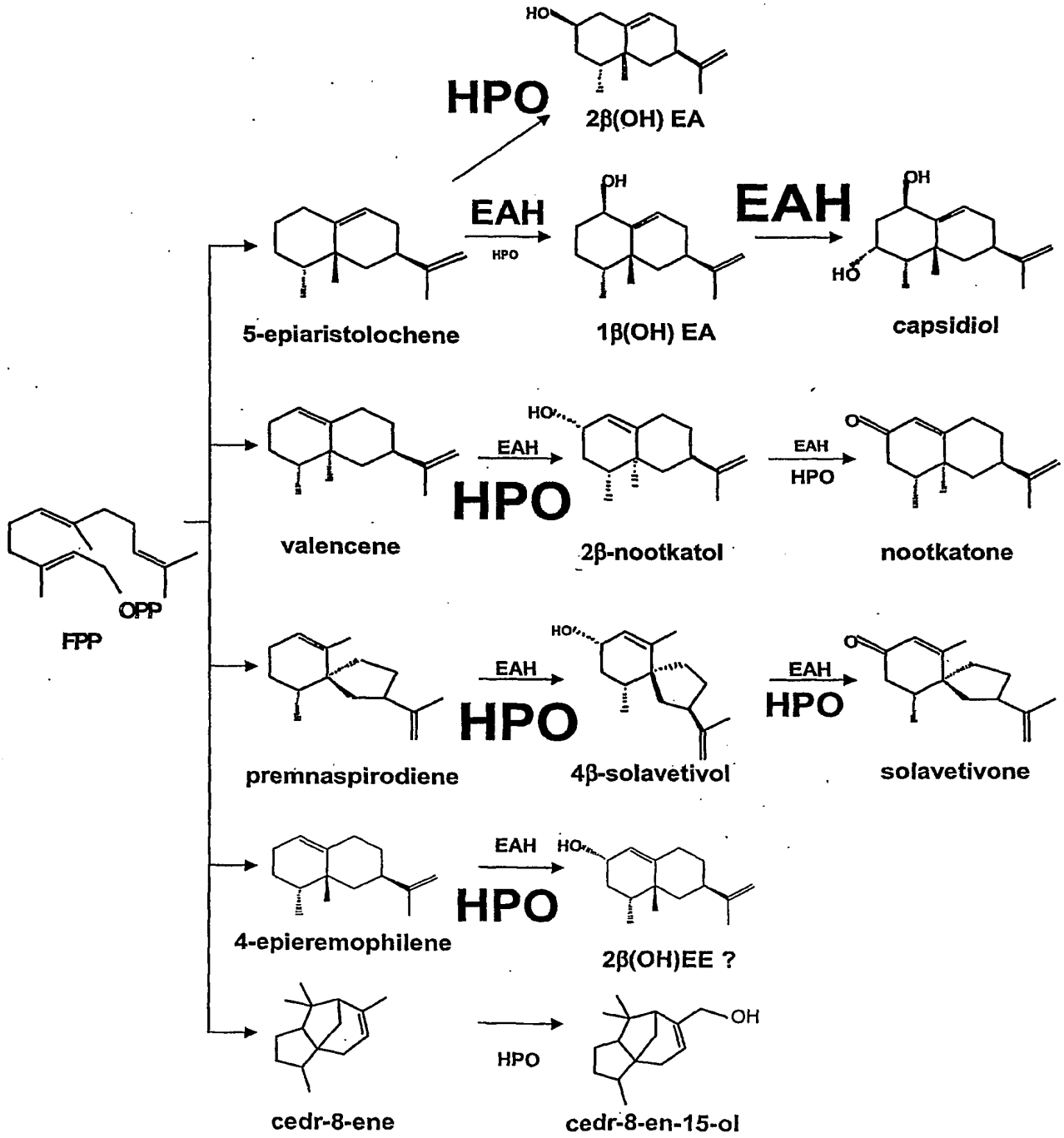


FIG. 9

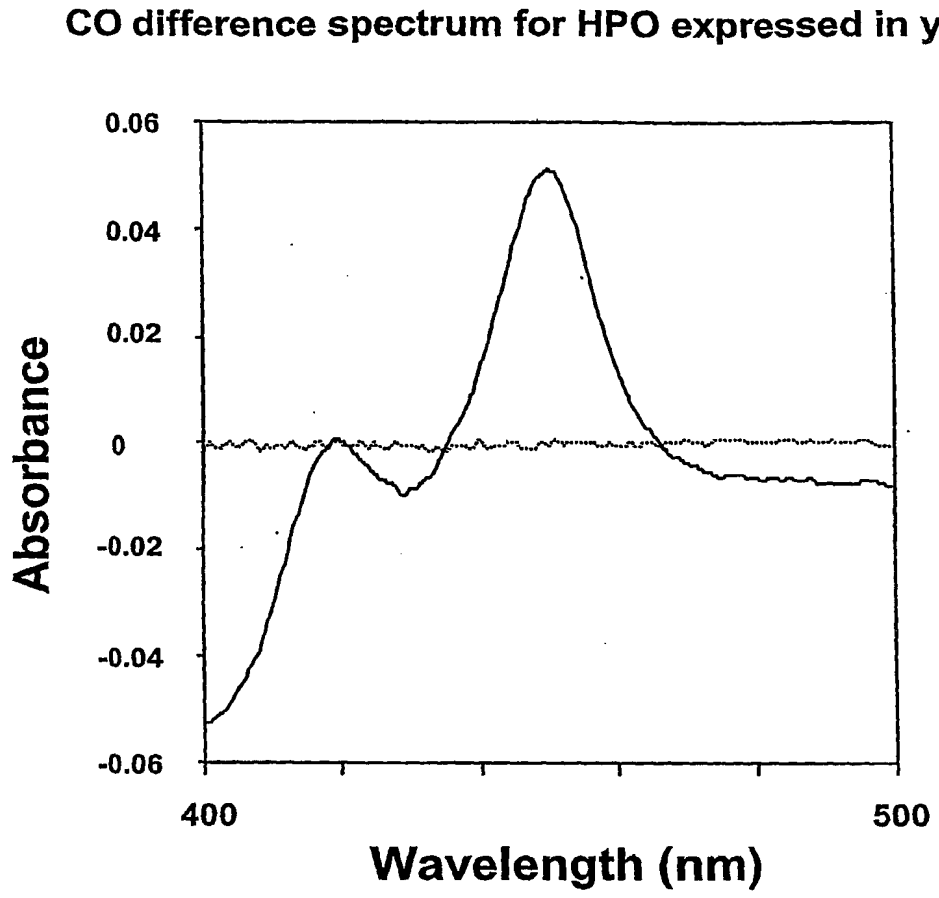
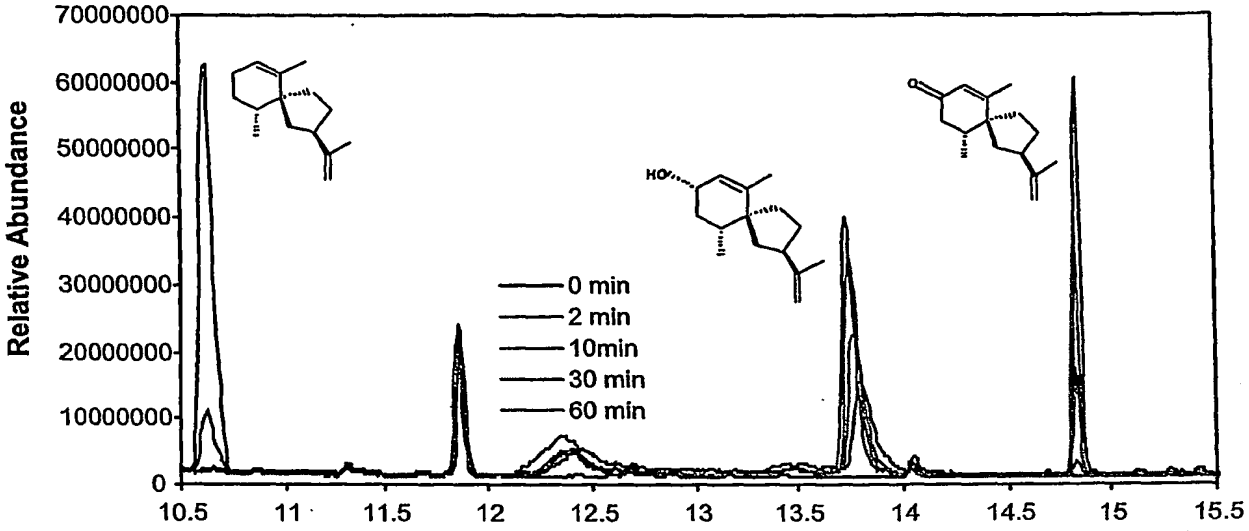

CO difference spectrum for HPO expressed in yeast
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
Absorbance
400
500
Wavelength (nm)

FIG. 10

1. Substrate: premnaspirodiene



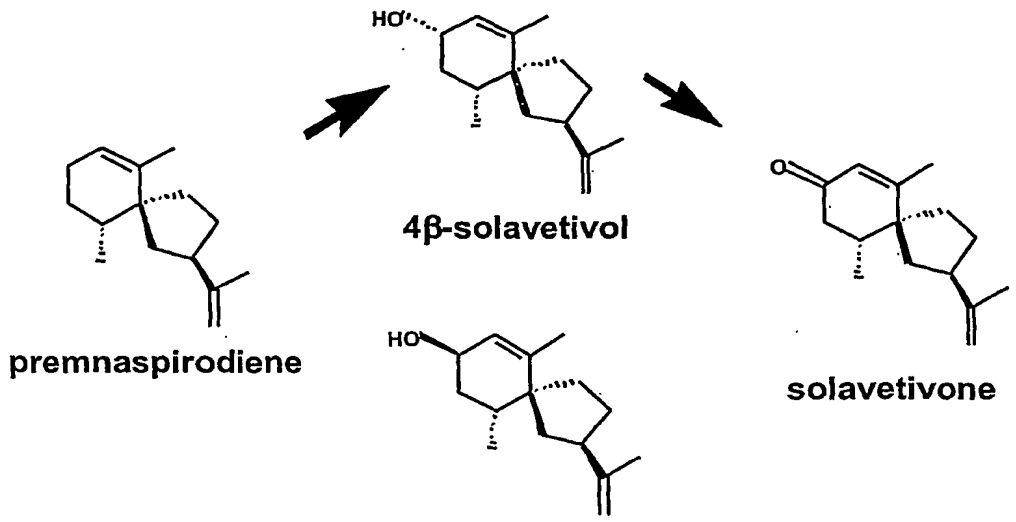
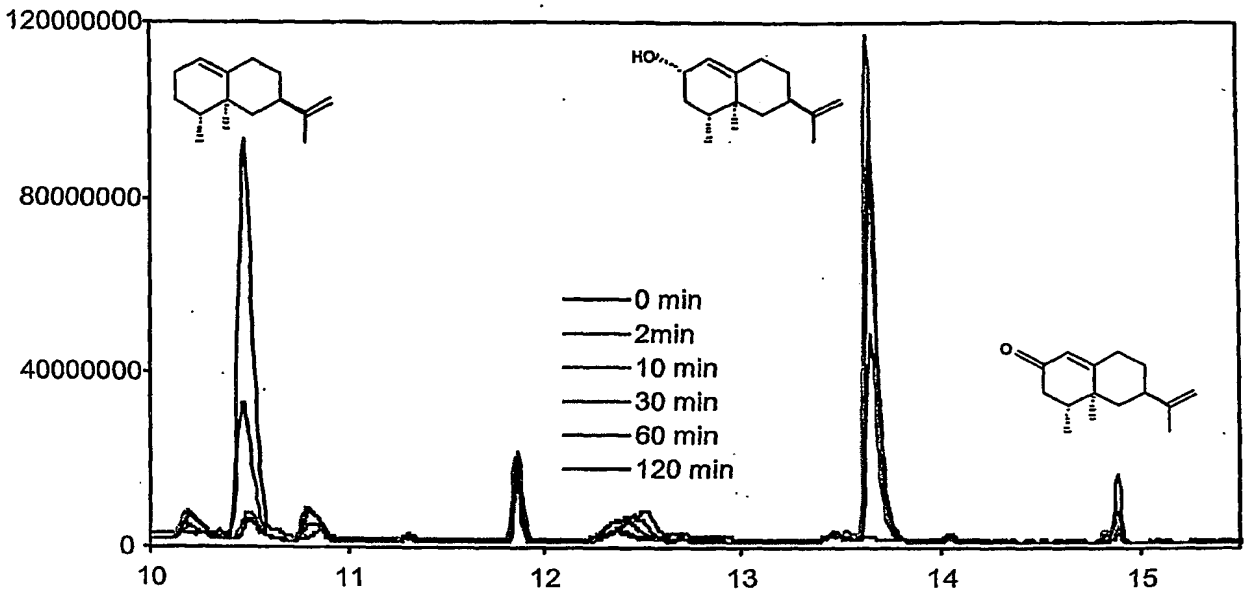
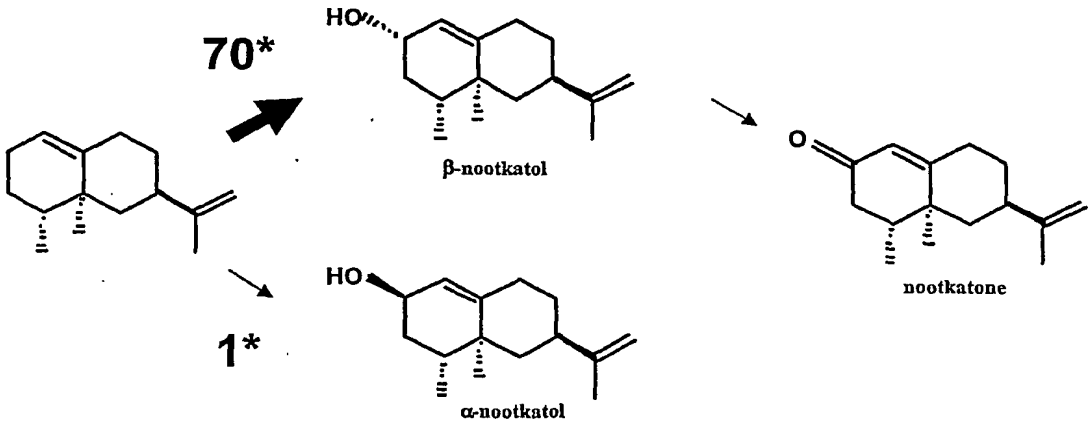


FIG. 11

Substrate: valencene



Summary of the reaction



* Relative specificity constant based on k_{cat}/K_m value

Fig. 12

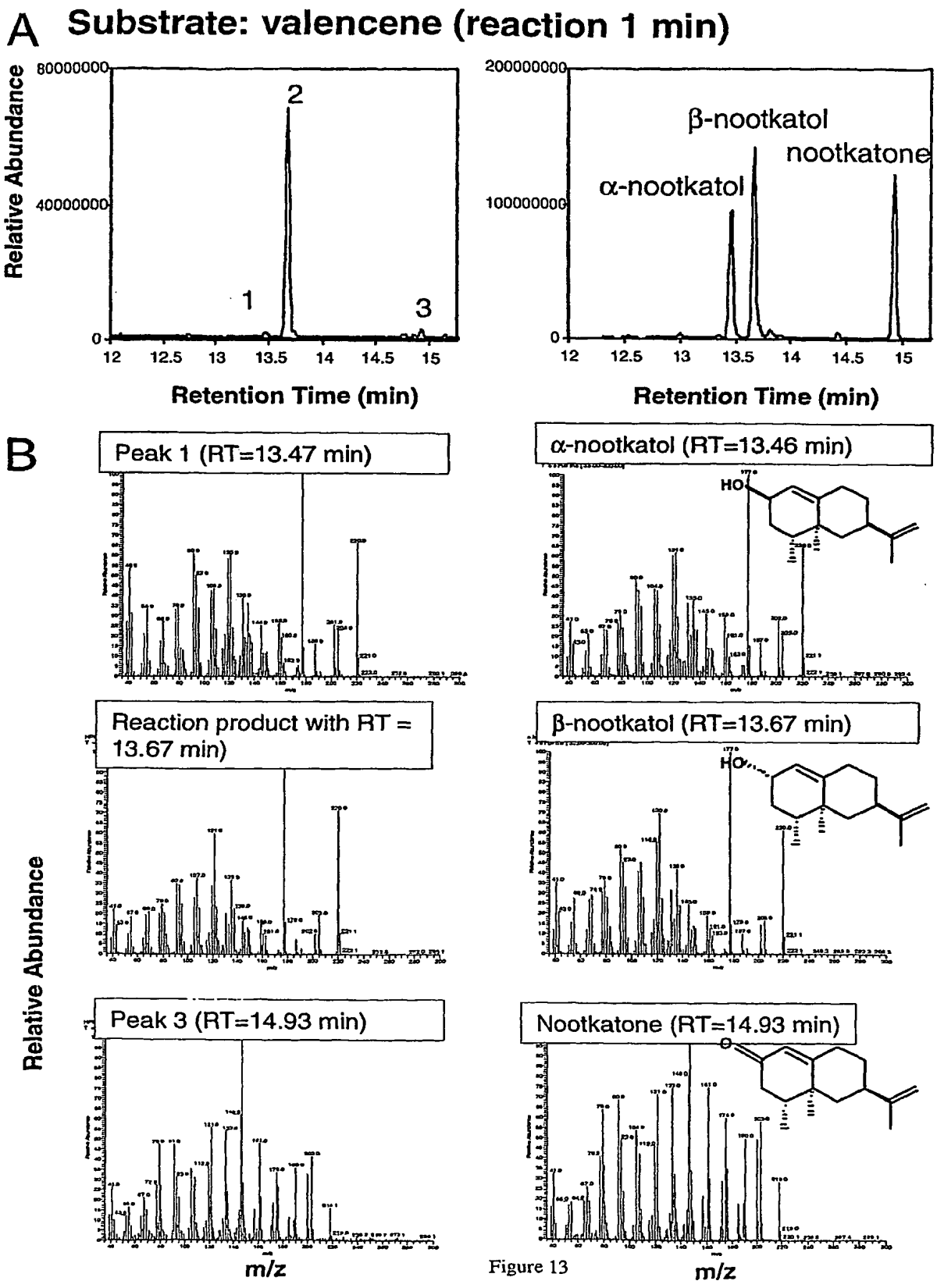
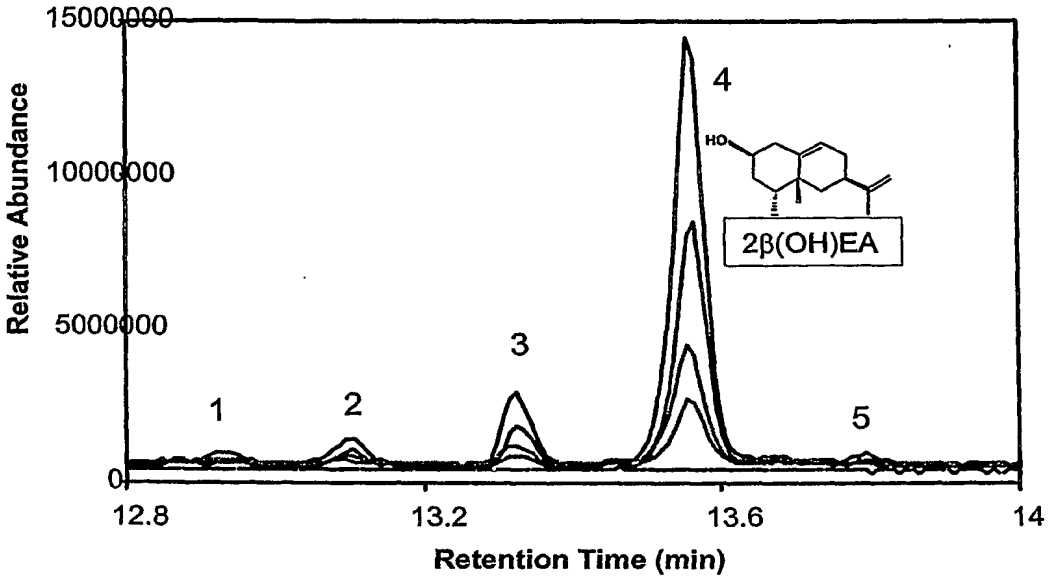

A
Substrate: valencene (reaction 1 min)
Relative Abundance
80000000
40000000
0
1
2
3
12
12.5
13
13.5
14
14.5
15
Retention Time (min)
200000000
100000000
β-nootkatol
nootkatone
α-nootkatol
Retention Time (min)
B
Peak 1 (RT=13.47 min)
α-nootkatol (RT=13.46 min)
HO
Reaction product with RT = 13.67 min)
β-nootkatol (RT=13.67 min)
Peak 3 (RT=14.93 min)
Nootkatone (RT=14.93 min)
Relative Abundance
m/z
m/z

Figure 13

Substrate: 5-epi-aristolochene (Reaction 0, 1, 2, 5, 10 min)



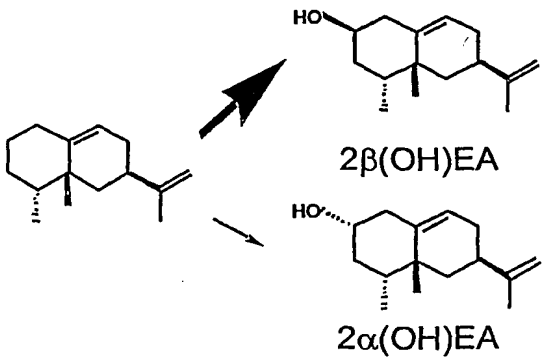


FIG. 4A

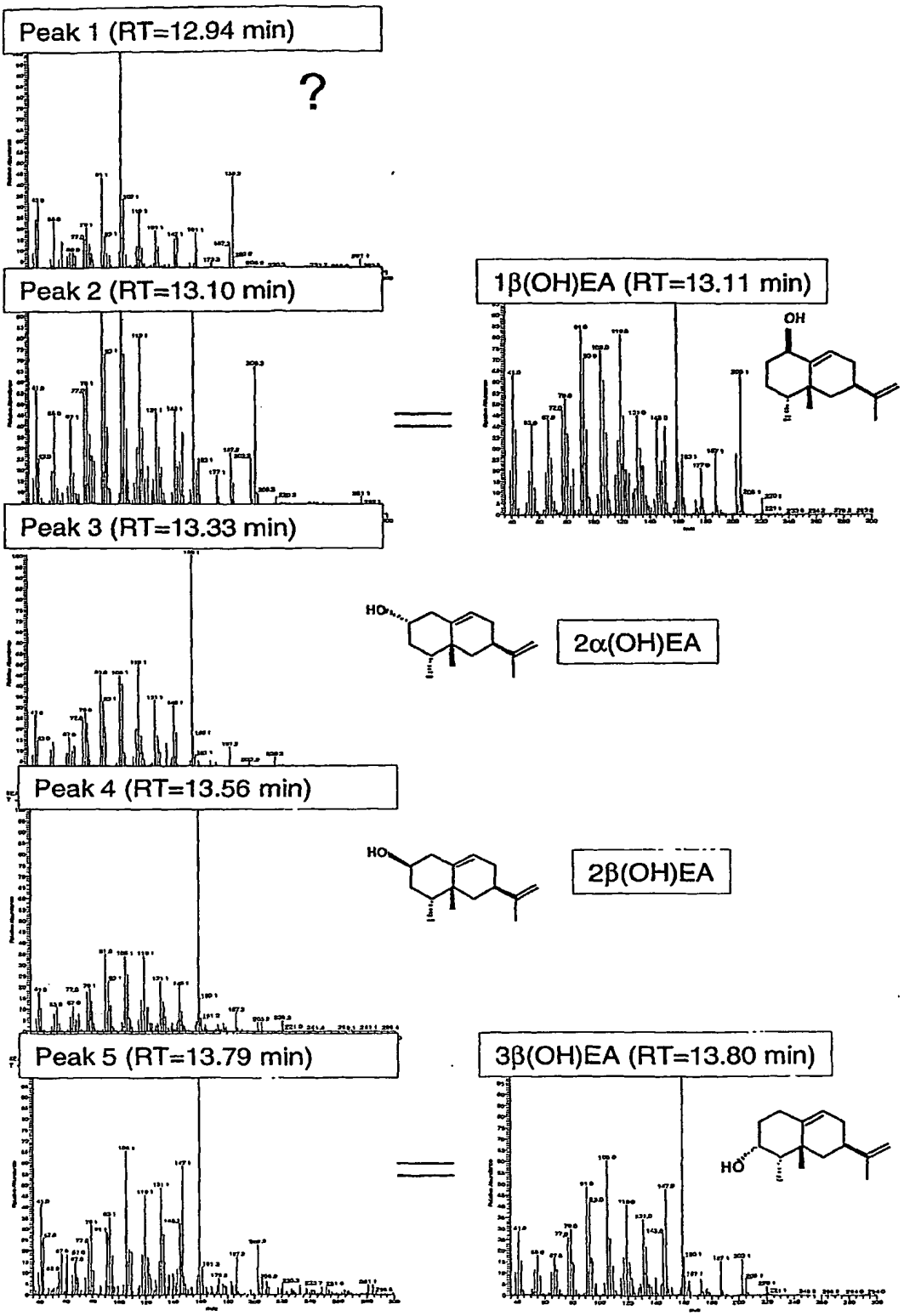
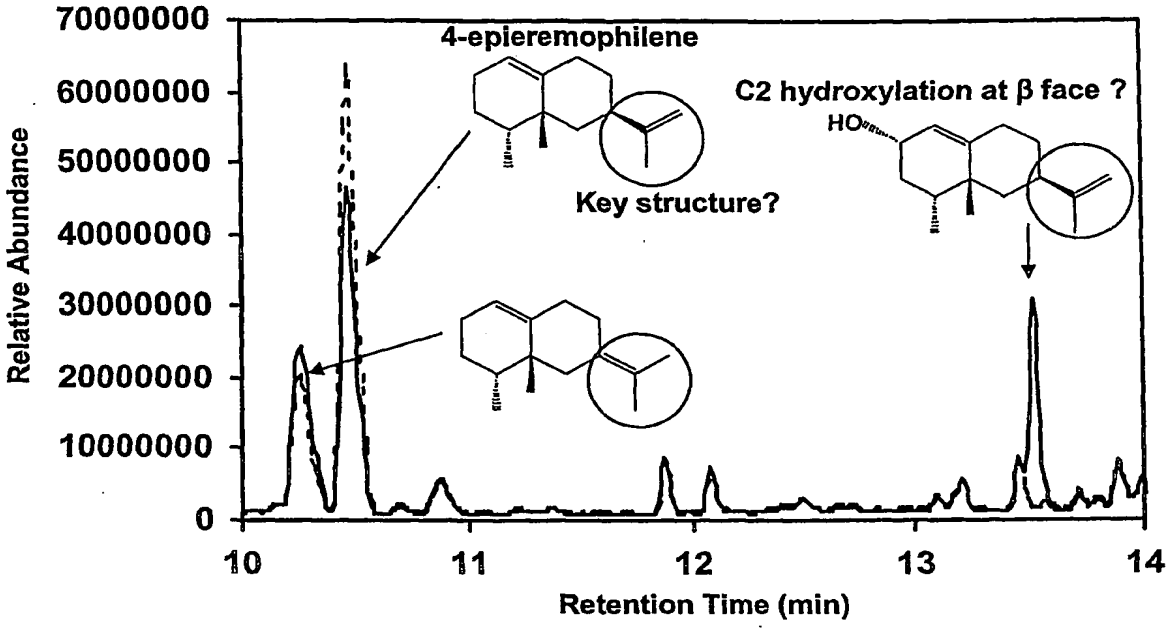


Figure 14 cont.

Substrate :4-epieremophilene+double-bond isomer (Reaction 5min)



Substrate : α-cedrene (Reaction 5 min)

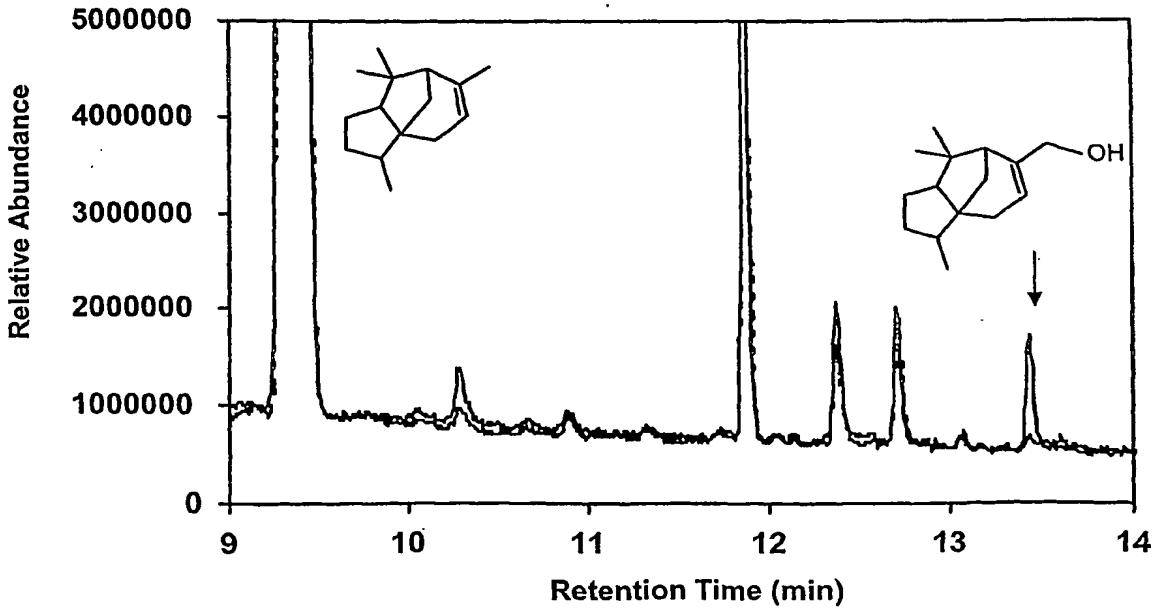


FIG. 15

Comparison of enzyme kinetics of HPO for various substrates relative to the previously characterized EAH (5-epi-aristolochene dihydroxylase) hydroxylases (Ralston et al.)

enzyme	substrate	Reaction product	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\ s^{-1}$)
EAH	5-epiaristolochene	capsidiol	19.18 ± 2.90	0.493 ± 0.039	0.0257
	1β(OH)EA	capsidiol	1.74 ± 0.08	0.582 ± 0.006	0.3344
HPO	premnaspirodiene	4β-solavetivol	14.01 ± 1.94	2.116 ± 0.099	0.1510
		solavetivone	1.70 ± 0.15	0.077 ± 0.001	0.0453
	4β-solavetivol	solavetivone	1.24 ± 0.11	0.097 ± 0.001	0.0782
	valencene	α-nootkatol	11.47 ± 1.91	0.041 ± 0.002	0.0036
		β-nootkatol	7.35 ±1.20	1.929 ± 0.067	0.2624
	5-epiaristolochene	2β(OH)EA	3.25 ± 0.34	0.223 ± 0.005	0.0686
	4-epieremophilene	2β(OH)EE?	7.77 ± 0.99	0.819 ± 0.027	0.1054
	cedr-8-ene	cedr-8-en-15-ol	26.45 ± 1.08	0.059 ± 0.001	0.0022

FIG. 16

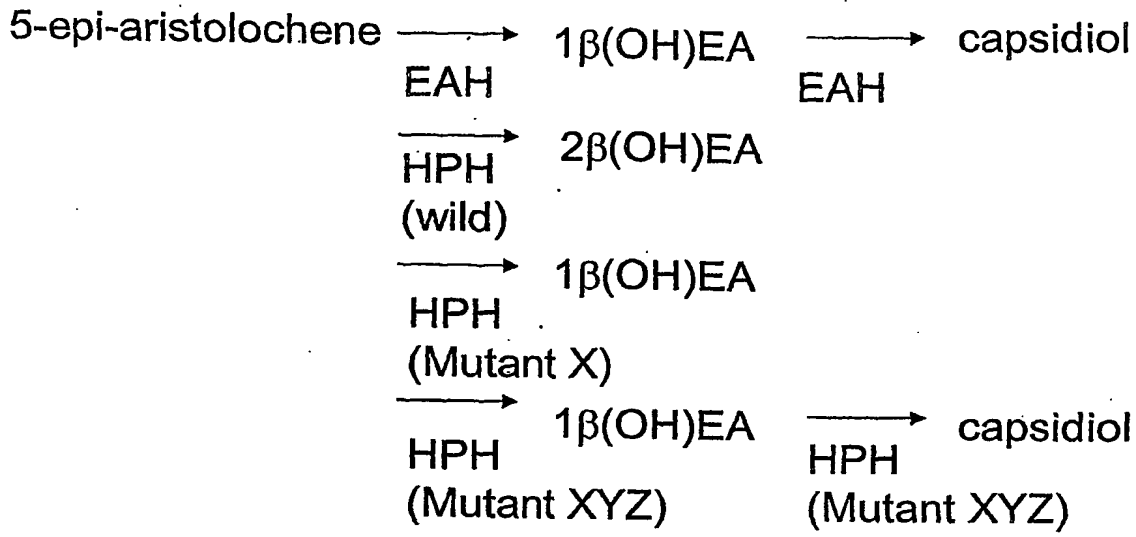
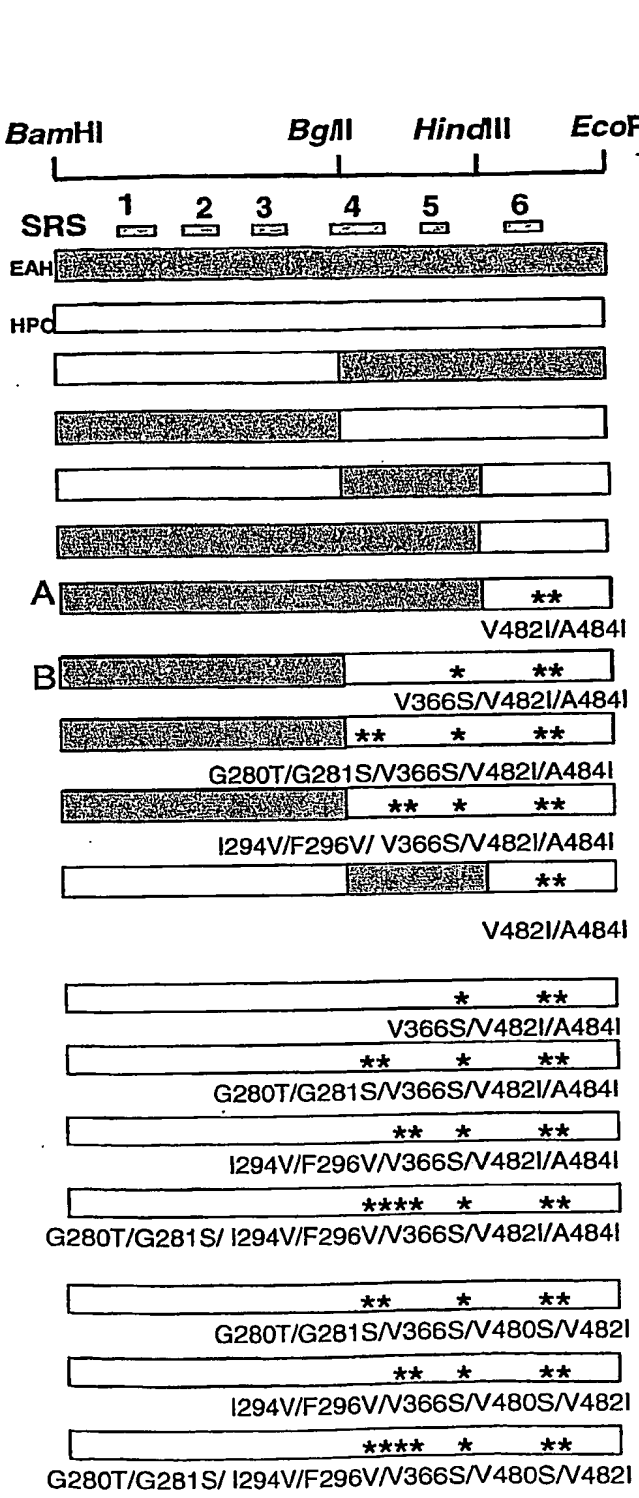

5-epi-aristolochene
EAH
1β(OH)EA
EAH
capsidiol
HPH
(wild)
2β(OH)EA
HPH
(Mutant X)
1β(OH)EA
HPH
(Mutant XYZ)
1β(OH)EA
HPH
(Mutant XYZ)
capsidiol

FIG. 17

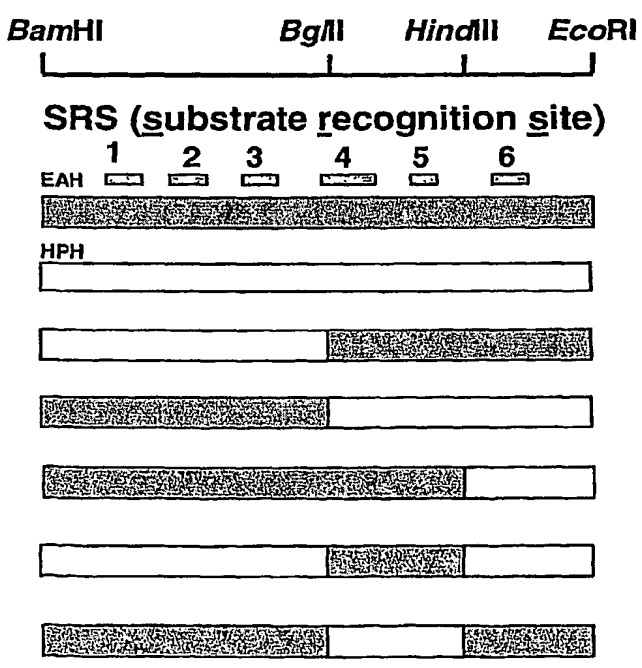


	EAH activity (nmol product/nmol enzyme/sec) Substrate: 5EA (40 μM)		Substrate: Iβ(OH)EA (25μM)
	Iβ(OH)EA	capsidiol	capsidiol
EAH	0.12	0.33	0.58
HPD	0.0073	0	0
	0	0	0
	0	0	0
	0.079	0	0
	0.005	0 Base peak 157	0
A	0.100	0.2521	0.3196
B	0.00518	0	0
	0	0	0
	0.1109	0.3628	0.8187
	0.06937	0	0.00864
	0.2138	0	0
	0.03169	0	0
	0.2275	0.04724	0.1137
	0.378	0	0.0521
	0.1594	0	0
	0.1791	0.01585	0.11701
	0.15632	0.00882	0.1015

Figure 18

Strategy

1. Domain-swapping mutation based on Substrate recognition sequences (SRS)
2. Reciprocal site-directed mutagenesis based on homology modeling with mammarlian P450s
3. Combination of domain-swapping and site-directed mutagenesis

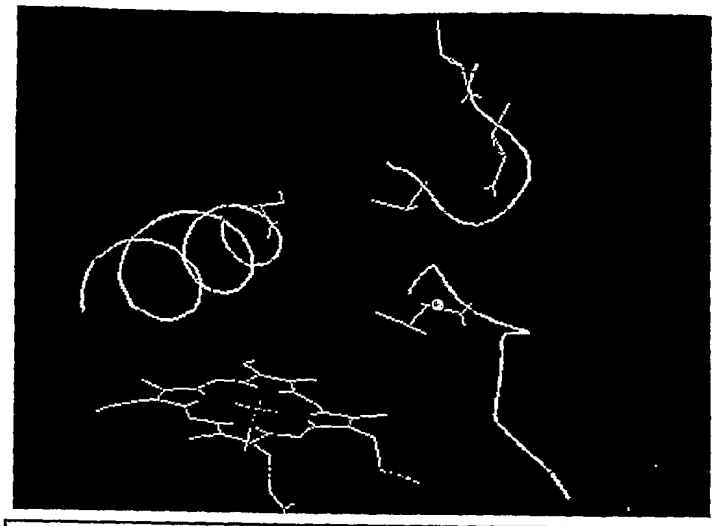


EAH activity
(nmol product/nmol enzyme/sec)

Substrate	5EA (40 μM)		Iβ(OH)EA (11μM)
Product	Iβ(OH)EA	capsidiol	capsidiol
	0.12	0.33	0.50
	0.0073	0	0
	0	0	0
	0	0	0
	0.005	0	0
	0.079	0	0
	0	0	0

Figure 19

2. Homology modeling and site-directed mutagenesis-mutants of HPO indicated by amino acid in native (wildtype) enzyme, amino acid position, then mutant amino acid.



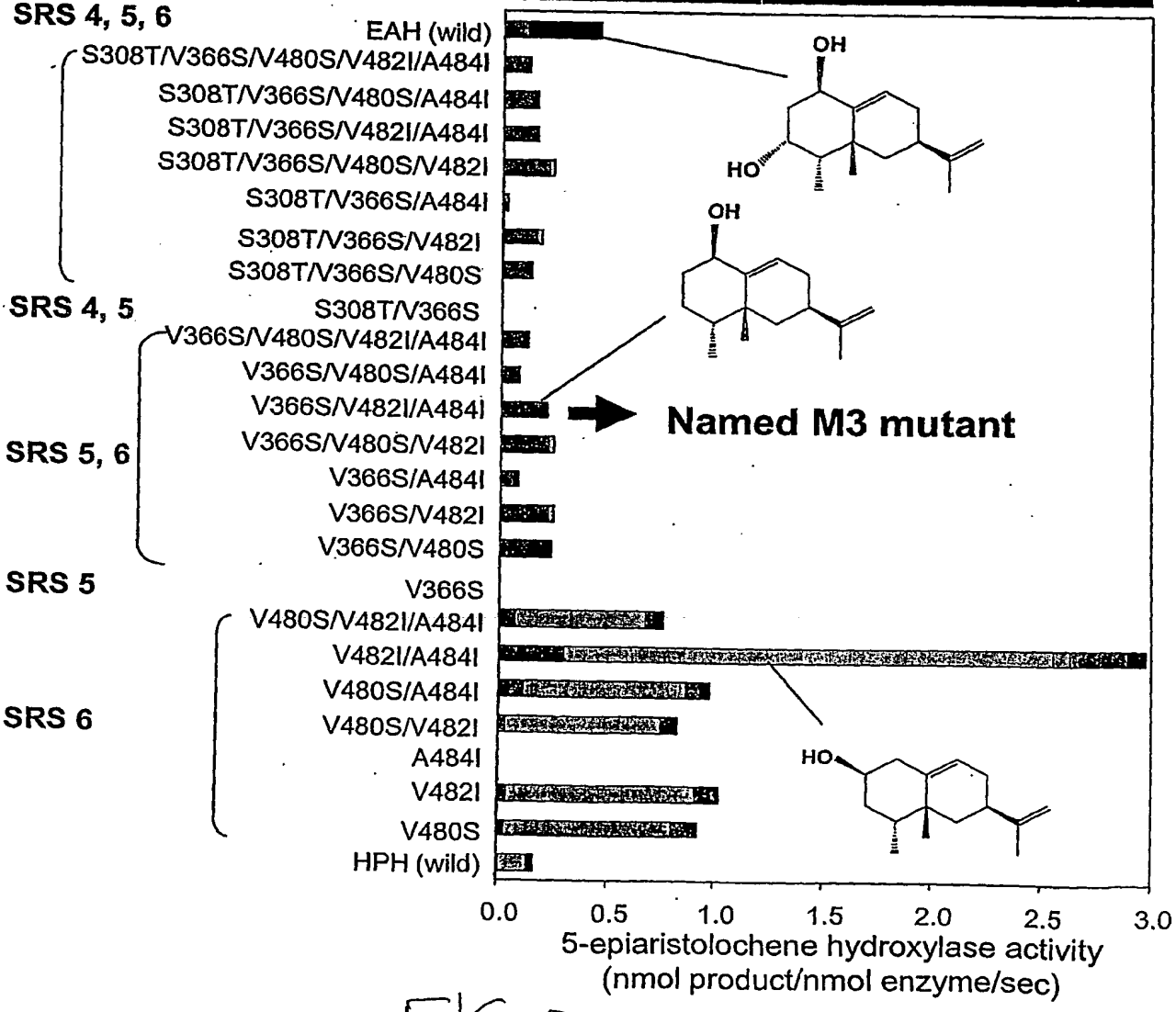
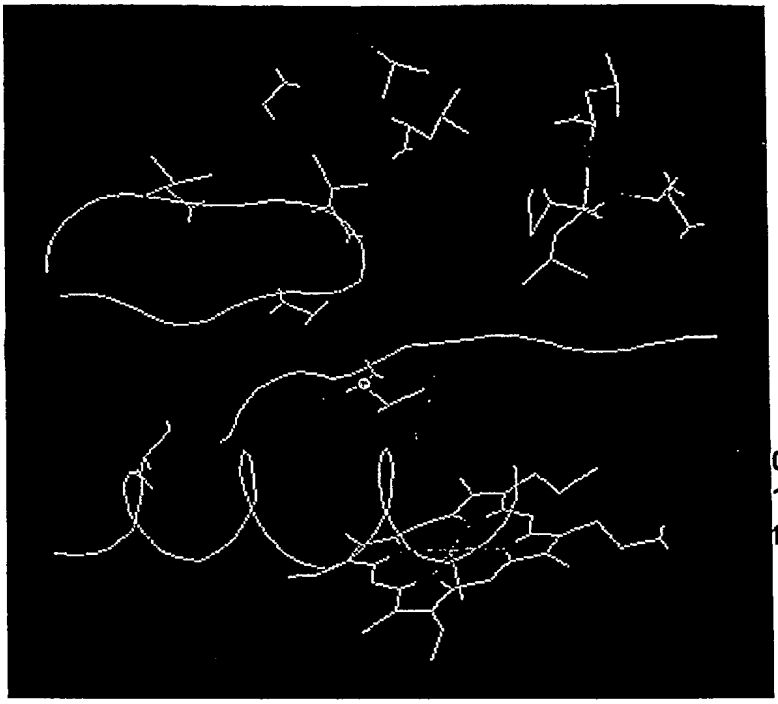


FIG. 20

2. Homology modeling and site-directed mutagenesis

SRS 1, 2, 4, 5, 6



09I
13Q
19S

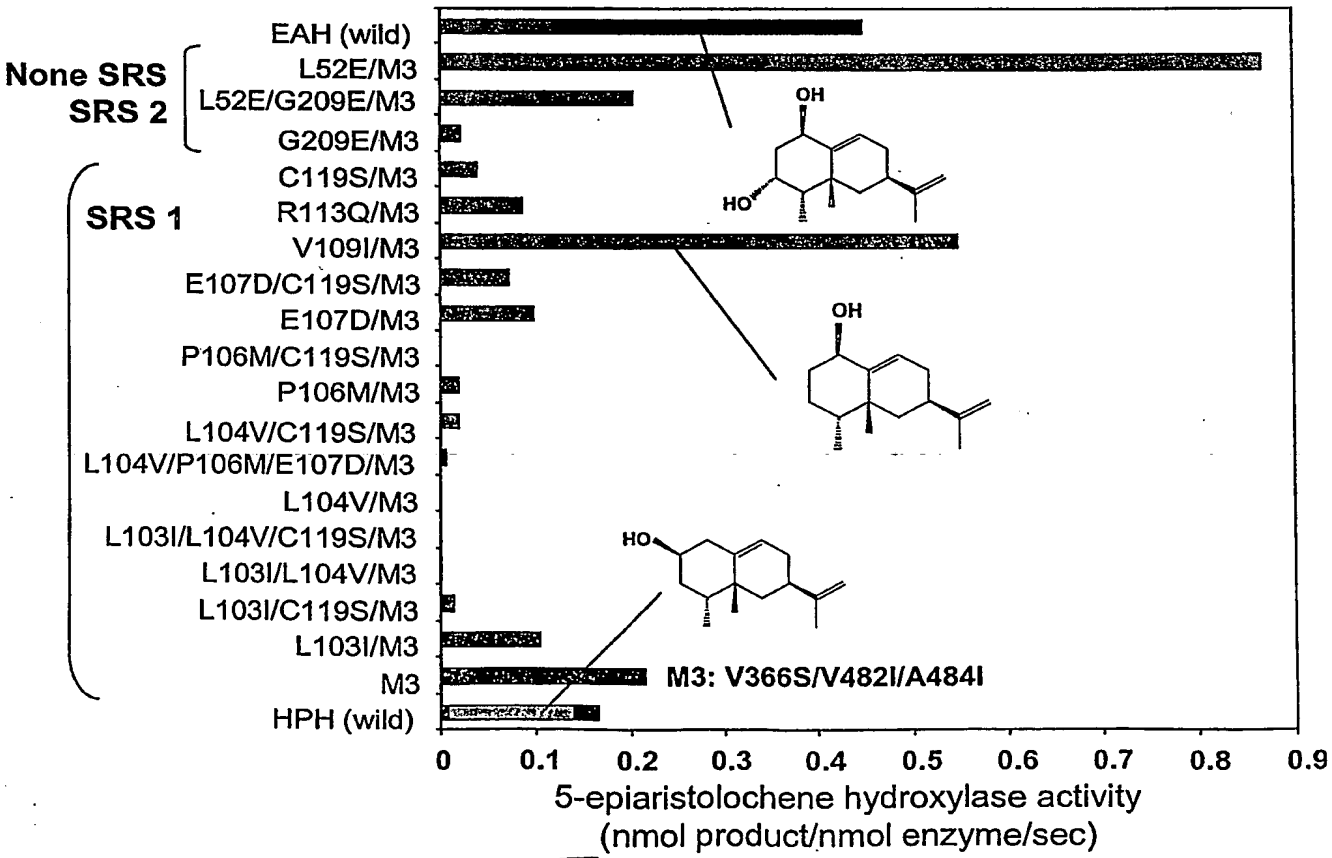
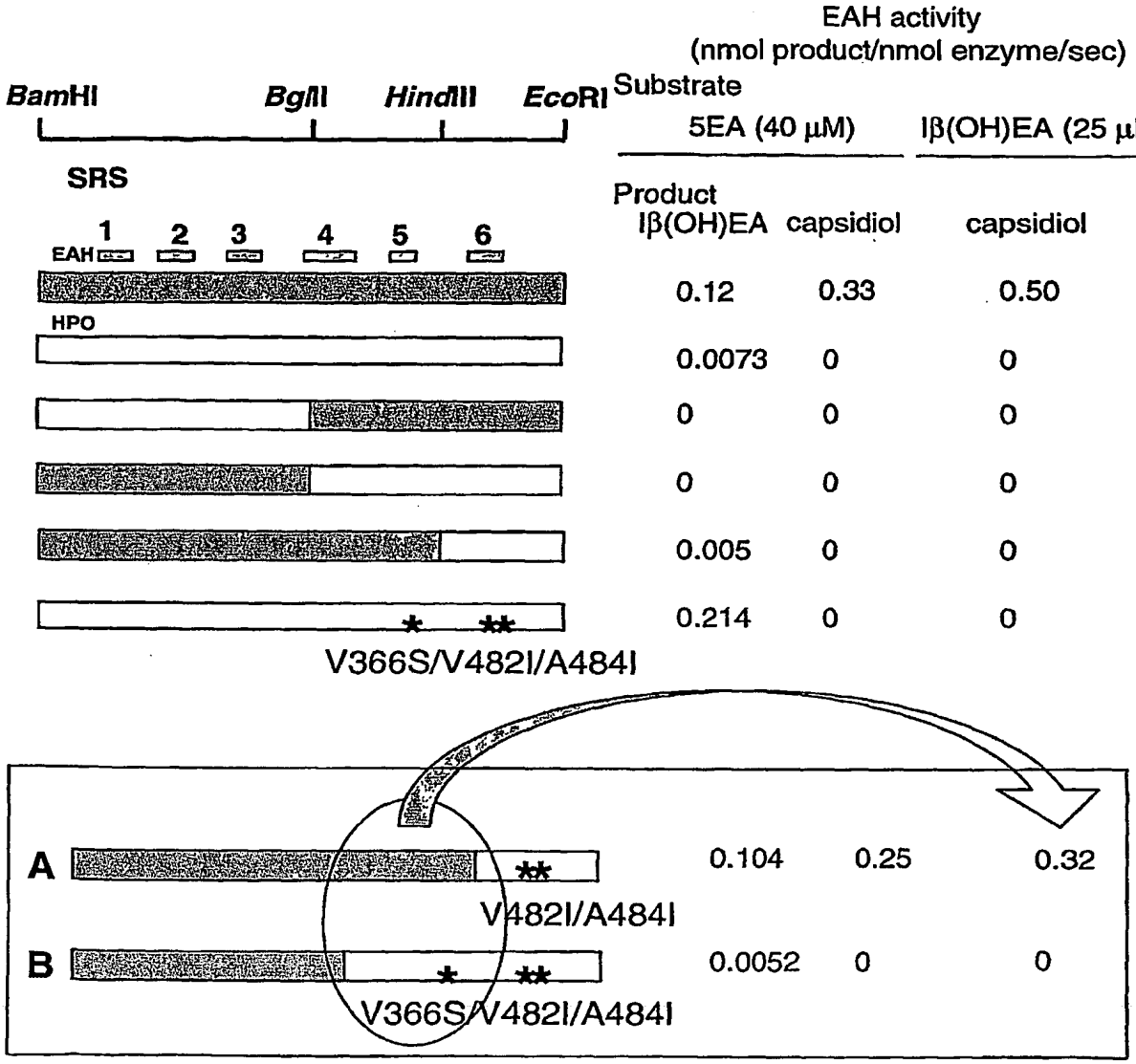


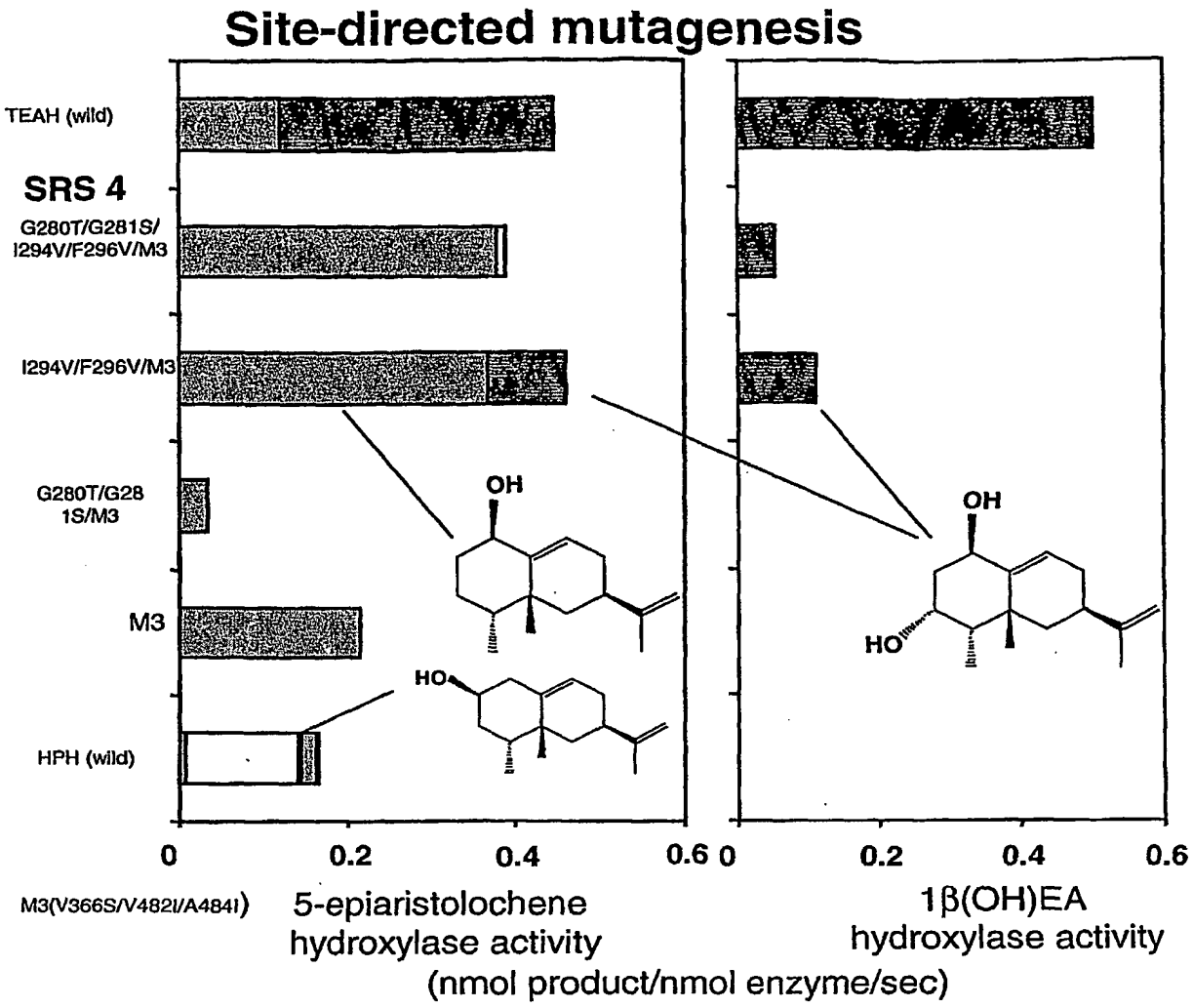
FIG. 21

3. Combination of Domain-swapping mutation and site-directed mutagenesis

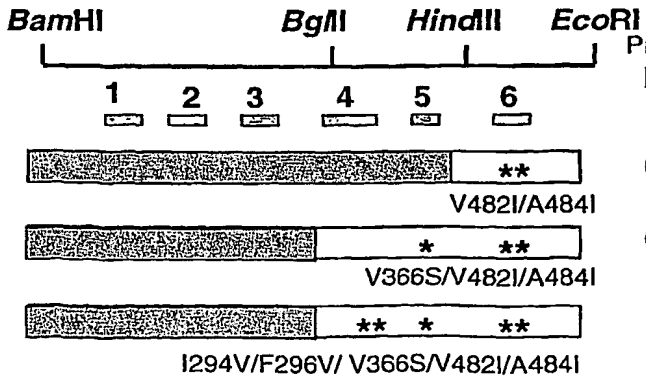


Key information on structural elements regulating second hydroxylation step

Figure 22



Domain-swapping and Site-directed mutagenesis



Substrate	EAH activity	
	5EA (40 μM)	Iβ(OH)EA (25 μM)
Product	Iβ(OH)EA capsidiol	capsidiol
V482I/A484I	0.100 0.2521	0.3196
V366S/V482I/A484I	0.00518 0	0
I294V/F296V/ V366S/V482I/A484I	0.1109 0.3628	0.8187

Figure 23